VOL. XV OCTOBER, 1931 No. 5

The Journal of APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

H. L. HOLLINGWORTH. Diurnal Variations in Suggestibility 431
ROBERT J. BALL. The General Emotionality of the Prisoner 436
ROBERT A. BROTEMARKLE. Determining the Points of Emphasis in College-Adult Level Testing 462
MILES A. TINKER AND DONALD G. PATERSON. Studies of Typographical Factors Influencing Speed of Reading. VIII: Variations in Color of Print and Background 471
MATTHEUS SMITH. Concerning the Magnitude of the Behavior Sample for the Study of Behavior Traits in Children 480
PAUL THOMAS YOUNG. Sex Differences in Handwriting 486
ERNEST R. HILGARD. A Study of "Warming-up" in the Case of a Task of More Complicated Perceptual-Motor Coordination 499
GEORGENE H. SEWARD. Students' Reactions to a First Course in Psychology 512
Notes and News 525
Book Reviews 529
New Books and Pamphlets 538

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio

Printed by
THE SCIENCE PRESS PRINTING COMPANY
LIME AND GREEN STS., LANCASTER, PA.

Entered as second-class matter at the post office at Lancaster, Pa., under the act of March 3, 1879.
Application for additional entry at Athens, Ohio, post office, pending.
Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1931, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University; W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston, H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Wooley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, gender and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00. Remittances and business communications should be addressed Lime and Green Streets, Lancaster, Pa., or to the Editor, James P. Porter, Ohio University, Athens, Ohio.

DIURNAL VARIATIONS IN SUGGESTIBILITY

H. L. HOLLINGWORTH

Columbia University

Suggestibility may be said to represent a partial dissociation. In dissociation either of two things may occur. An available cue may not articulate in effectiveness with other cues operative at the moment. Or an effective cue may operate unwittingly, that is, without its presence and potency being reported. Suggestion represents dissociation in the latter aspect.

A favorite account of drowsiness and sleep is in terms of dissociation. Since sleeping in a general way comes on as the day proceeds, we might expect that tendencies to dissociation would be more marked toward the end than at the beginning of the day's work. And since suggestibility is an aspect of dissociation, we might expect differences in suggestibility at different times of day. Theoretically, an individual should become more suggestible as he became more sleepy. Suggestibility should increase during the course of the day, at least in a roughly progressive way.

We waive the question of general suggestibility, and also the problem of the various forms of suggestion—such as personal and impersonal suggestion. For the present purpose we mean by suggestibility only the particular phenomenon represented in the situation here tested—that of the size-weight illusion.

Any perceptual illusion may be said to constitute a case of suggestion. The usual rule is that report is unwittingly determined by some incidental cue. This cue is extraneous to the explicit field on which the report is made. Suggestion, in the impersonal sense at least, is illustrated by the potency of such a collateral determinant.

The size-weight illusion is a classical instance. As the explanation usually runs, the visual size of the larger block in-

stigates lifting movements more vigorous than those unwittingly adopted for the smaller block. Relative difference in visual size determines differences in innervation. The resulting apparent difference in reported heaviness is due to this unreported difference in innervation, and its consequences in the behavior of the lifted weight.

Fifteen subjects, ten men and five women, were given the size-weight illusion test five times daily for 26 days. The test hours daily were at or about 8:00 and 10:30 A. M. and 12:30, 3:30, and 6:00 P. M., varying slightly with the subject, since all were tested in succession at intervals of about one and one-half hours.

The standard (large) block weighed 55 grams. From the array of smaller blocks one was, at each test, selected as equal to the standard. The comparison weights differed by increments of 5 grams, ranging from 15 to 80 grams. One selection was made at each test, the subject being allowed to make as many comparisons as were needed to find the most satisfactory equivalent of the standard.

During the first 16 days definite changes occurred as the result of the repetition of the test. These changes will be elsewhere reported in detail, but their general character was shown by a progressive increase in the magnitude of the illusion. After 16 days the effect of repetition had ceased and the curve of error (in choosing too light a comparison weight) from day to day became a straight line. Practise having thus been eliminated, a study could be made of changes from hour to hour within a given day. The last ten days (17th to 26th) were used for these results.

The data here used therefore represent five daily tests, on each of ten days, of 15 subjects, after 16 preliminary days of practise. They indicate certain changes in the magnitude of the illusion in the case of the size-weight block comparisons.

The following table gives, for each of the 15 subjects, the average weight chosen at each test hour on the 10 days, with the Average Deviation of each average. At the foot of the table are given for each test hour the average weight chosen

(by the fifteen subjects) and the Probable Error of this average. Subtracting the weight chosen from 55 grams (the objective weight of the standard) gives the magnitude of the illusion error.

Average weight chosen at different times of day on 10 days, following 16 preliminary practise days

SUBJECT	APPROXIMATE HOURS OF THE DAY OF EACH TEST									
	A.M. 8:00		10:30		P.M. 12:30		3:30		6:00	
	Aver.	A.D.	Aver.	A.D.	Aver.	A.D.	Aver.	A.D.	Aver.	A.D.
1	25.5	1.8	24.0	2.4	24.0	1.6	24.5	1.8	24.5	0.9
3	20.0	0.0	20.0	0.0	20.0	0.0	20.0	0.0	20.5	0.9
4	33.5	3.5	31.5	2.6	32.0	2.4	32.0	2.4	30.5	3.6
5	19.5	0.9	20.5	0.9	20.5	0.9	20.0	0.0	19.0	1.6
6	15.5	0.9	16.5	2.1	15.5	0.9	15.0	0.0	15.5	0.9
7	24.0	1.6	25.0	0.0	24.5	0.9	24.5	0.9	25.0	0.0
8	25.0	0.0	25.0	0.0	25.0	0.0	25.0	0.0	25.0	0.0
9	15.5	0.9	15.0	0.0	15.0	0.0	15.0	0.0	15.5	0.9
10	24.5	0.9	24.0	0.9	25.0	0.0	25.0	0.0	25.0	0.0
11	26.5	2.7	25.5	2.2	25.5	1.8	25.5	1.8	27.0	3.0
12	25.0	0.0	25.0	0.0	25.0	0.0	24.5	0.9	25.0	0.0
13	23.5	2.1	25.0	2.0	25.5	0.9	25.0	1.0	26.0	1.6
14	23.5	2.6	22.0	2.4	20.0	0.0	21.0	1.6	21.0	1.6
15	29.0	1.6	26.0	1.6	26.5	2.3	26.5	2.1	26.5	1.5
16	34.0	0.9	36.0	1.8	34.5	0.9	33.5	2.5	35.5	0.9
Average Weight	24.3		24.1		23.9		23.8		23.4	
P.E. of Average	.76		.70		.74		.74		.82	
Average Illusion Error	30.7		30.9		31.1		31.2		31.6	

Although the differences are small the average weight chosen becomes gradually smaller and smaller in the course of the day, decreasing by 1 gram from the early morning to the late afternoon trial. The error, that is, the magnitude of the illusion, gradually increases from an average of 30.7 to an average of 31.6 grams. This increase of about one gram over the morning illusion of about 30 grams represents an increase

of about 3 per cent. In this form of suggestion these subjects may therefore be said to be about 3 per cent more suggestible at the last trial than at the first. The greatest step in this gradual increase occurs at the late afternoon trial.

Perhaps the comparison with greatest statistical validity is that of the evening with the morning hours, omitting the intermediate noon-day test. This comparison gives the following set of averages, each figure for each subject being the average of 20 tests. The final averages are thus the result of 300 morning and 300 evening trials.

Average weights chosen in morning and evening trials, each figure the average of 20 tests

SUBJECT	MORNING AVERAGE	EVENING AVERAGE
1	24.8	24.5
3	20.0	19.8
4	32.5	31.2
5	20.0	19.5
6	16.0	15.2
7	24.5	24.5
8	25.0	25.0
9	15.3	15.3
10	24.5	25.0
11	26.0	26.2
12	25.0	24.7
13	24.2	25.5
14	22.8	21.0
15	27.5	26.5
16	35.0	34.5
Average Weight	24.2	23.9
Average Illusion	30.8	31.1

The difference between the average morning and the average evening illusion is small. It amounts to .3 grams, the P. E. of this difference being .9, larger, that is, than the difference itself. But the difference is in the direction of greater evening suggestibility. In the case of 9 of the 15 subjects the difference is in this direction, and in 3 cases there is no difference. In only 3 of the 15 cases, that is, is the direction of the difference opposed to the general direction.

The results as a whole are not striking in magnitude, but their direction as shown in all the comparisons is consistent with the theoretical implication. In so far as our results are concerned the indication is that when a difference is present evening is a more suggestible period than morning. The change from early forenoon to late afternoon is in this case an increase of 3 per cent.

A change in the variable error of such a judgment might be supposed to be due to waning interest and consequent carelessness as the day's work proceeded. But we are dealing here with a constant error. There is no reason why mere carelessness should increase rather than decrease the illusion. The result seems instead to be due to a positive increase in the strength of the illusion motive or else to increased susceptibility to that motive. Since the objective features of the situation remained unchanged, the explanation in terms of increased susceptibility seems to be the more plausible one.

Such measures of changes in suggestibility with time of day have rather interesting results for applied psychology. If individuals are on the whole more open to suggestion in the evening than in the morning, psychotherapy through waking suggestion and persuasion may well take advantage of the fact. The observation is indeed entirely in accord with the practise of first putting the patient into a somnolent state before proceeding with verbal therapy. If people are more suggestible when they read the evening paper than when scanning the morning journal, this fact has a certain interest also for advertising, for propaganda, for educational publicity in general. Furthermore, the indications are that individual differences exist along with the average tendency. These individual differences themselves it may perhaps be useful to take into account in practise.

It must, however, be noted that such diurnal changes as those found in the present experiment are so slight and so unreliable that the results may be said only to raise the question, not to settle the issue.

THE GENERAL EMOTIONALITY OF THE PRISONER

ROBERT JAUDON BALL
Berkeley, California

INTRODUCTION

Recently the writer has undertaken a study of the motor learning processes of unstable and neurotic men. The Penitentiary at San Quentin, California, was chosen as a suitable place where a selection of emotionally unstable individuals might be found in relatively large numbers. In order to facilitate this selection some definite instrument had to be used as it would have been a much slower and uncertain process to have depended upon subjective diagnoses made by prison officials, physicians, etc. The instrument used for this purpose was the original Woodworth Personal Data Sheet described by Franz (1), and a modification of the same devised by Cady (2). The latter had to be modified slightly for our needs; *i.e.*, changing a few questions from present to past tense, but was discarded after 43 cases had been run. The original Personal Data Sheet of Woodworth was substituted and given to 92 men. The total of 135 cases was necessary in order that the required 20 neurotics and emotionally unstable men could be selected for the motor learning tests.

There has been some objection to the use of the Woodworth Personal Data Questionnaire, especially the Matthews' revision (4), by Flemming and Flemming (5) who attack the validity of the measure. The criterion for the attack is based upon the scores obtained on a rating scale used in an investigation by C. W. Flemming (6) in 1922 in an attempt to analyze those factors other than "general intelligence" which influence school success and achievement. The rating scale used in part of the original investigation was described and used by Terman (7) in the compilation of data used in his

"Genetic Studies of Genius." Out of eleven traits described in this five-point rating scale, one of them (No. 6) concerns itself with emotional balance. The scores attained on this particular trait (range 1–9) were correlated with scores made on the Matthews' Revision of the Woodworth Questionnaire, the latter being given at a date *two years* after the teachers' ratings were made. The records of 88 girls attending the Horace Mann High School in New York were available for correlation. The result, as might be expected, was not satisfactory. The correlation between these teachers' estimates of emotional balance and the Matthews' Revision of the Woodworth Questionnaire was minus (–) .1481. However, this proves nothing, as rating scales even though used by experienced raters can not be considered satisfactory criteria against which an entirely different type of test can be correlated. The Woodworth Questionnaire is in itself a type of self-rating scale and a subjective evaluation of the subject's own personality. The questionnaire is nothing more than a list of symptoms commonly present in various types of instabilities, neuroses, etc., and the subject is required to answer by "yes" or "no" the presence or absence of these symptoms. Can we then say that the subject's own analysis of himself can be adequately correlated with a teacher's estimate of his emotionality made two years previously?

Hollingworth (3) has definitely shown that the total emotionality of an individual is in most cases totally altered after the causal stress producing an abnormal state is removed. He refers here to neurotic symptoms appearing during the war and the disappearance of same following the armistice. The median scores on the Personal Data Sheet of a pre-armistice group of 117 soldiers was 20–25 points, whereas the median scores of a post-armistice group of 202 soldiers was but 10–15 points. Is it not possible that the emotionality of persons in civil life and school children can also be altered by the removal of motivating causes? This would be especially true in many instances during the adolescent period of

a child's life when the total organism is subject to many strained situations made up of both biologic and environmental factors. During the space of two years many changes take place in the young adolescent's make-up. What was true at the first observation is not necessarily true on the second observation two years later. Of course our measures are not going to be valid when they are compared on this basis. It seems as though we are trying to put our observations in psychology on a basis with the mathematical accuracy of physics and related sciences before we have even isolated the elements to be measured! This is especially true in this instance where we are dealing with the most elusive and subtle qualities of the human personality, the emotions. To attempt to treat such an obscure and vague subject as the emotions with the mathematical exactness of statistical procedure adopted for application to tangible entities, seems nothing more than characteristic of a certain lack of perspective. Psychology deals with a very vacillating and uncertain organism, and certain branches of its observations, such as described above, can not be interpreted with mathematical exactness until the observations themselves become more exact.

Another fact might also be borne in mind. The section pertaining to emotional balance in the rating scale used by Flemming is most inadequate to give any teacher a clear concept of what the term includes. The degrees of stability can not be satisfactorily covered in any five-step rating scale on one line. Such a scale would of necessity be a complete emotional inventory covering much space and a very thorough understanding and knowledge of the individual concerned by the rater.

What would constitute an adequate test of validity of such a questionnaire as the Woodworth Personal Data Sheet would not be a simple rating scale used as criterion, such as described above, but a test more far reaching in its content and absolutely devoid of subjective judgments. Such tests have not as yet been devised to any degree of accuracy. The writer

has been developing such a test (8) but at the present writing all the data are not complete. Other types of tests and measures should be devised, and when any one of these, or all of them can reliably select so-called unstable individuals from groups of well diagnosed cases of psychotics, neurotics, psychopaths, hysterics, schizophrenics, etc., *then* the particular measure or measures might be considered a criterion against which an emotional inventory questionnaire could be correlated. As it now stands, we must rely solely upon subjective judgments of well trained observers, or upon the self-analysis of individuals under consideration who are requested to answer questionnaires or similar instruments of selection.

PROCEDURE

The nature of the investigation was such that the age range was limited to a span of 9 years; *i.e.* from 17 to 25 years of age inclusive. Consequently no information on the older prison group was obtained. The first group of 43 men selected were those who drifted into the prison hospital and included a large percentage of chronic complainers. These were the types originally sought, but as the process was slow and uncertain another method was later adopted by which men in the respective age groups were selected from the prison files and tested in groups of ten. No foreign born or illiterates were included. The tests were conducted in one of the hospital dormitories which was most adequate and free from any disturbing factors.

The first group of 43 men was run on the Cady modification of the Woodworth Questionnaire which was later abandoned as too youthful and the original Woodworth Questionnaire was substituted. Table I shows the age distribution of the entire group. This table includes 137 cases, 2 of which were rejected because of failure to complete the tests and are not included in the questionnaire groups tabulated below.

It was necessary to test this number of individuals in order to select twenty men who showed definite neurotic symptoms.

TABLE I
Age distribution of the group surveyed

AGE	FREQUENCY
25–6 to 25–11	1
25–0 to 25– 5	0
24–6 to 24–11	4
24–0 to 24– 5	7
23–6 to 23–11	4
23–0 to 23– 5	5
22–6 to 22–11	3
22–0 to 22– 5	10
21–6 to 21–11	15
21–0 to 21– 5	16
20–6 to 20–11	23
20–0 to 20– 5	23
19–6 to 19–11	11
19–0 to 19– 5	11
18–6 to 18–11	0
18–0 to 18– 5	3
17–6 to 17–11	1
Total	137

Mean	20 years 11½ months
S.D.	1.56
P.E.	.092

The criterion set was a minimum of 30 unfavorable responses. According to Franz (1) "probably any individual who answers 20 of the questions wrongly should be suspected of instability. . . . If the number of 'wrong' answers is greater than 30 grave suspicion of abnormality is warranted." In the Cady modification 85 questions are included, but 12 of them are so-called "dummy" questions designed to cast off any suspicion as to the nature of the test. Therefore only 73 of the questions are significant, and the set criterion of 30 unfavorable responses comprises 41.1 per cent of the entire test. On the other hand, the original Woodworth Questionnaire contains 116 questions and no "dummies." The criterion of 30 unfavorable responses in this instance comprises

25.9 per cent of the entire test. Due to the fact that the Cady modification was designed for juvenile groups, and consequently not as significant for an adult group, the same criterion was set for both forms.

RESULTS

Table II shows the distribution of the 43 scores on the Cady Modification. It will be noted that the mean score is 20.03 unfavorable responses. This number comprises 27.4 per cent of all the questions. The Standard Deviation of the distribution is 9.35 with a P.E. of $\pm$ 0.96. In this distribution only six individuals score above the set criterion of 30 unfavorable responses. This number comprises approximately 14 per cent of the entire distribution for this test.

TABLE II
Distribution of scores on the Woodworth-Cady Questionnaire

SCORE	FREQUENCY
50–54	1
45–49	0
40–44	0
35–39	1
30–34	4
25–29	6
20–24	10
15–19	8
10–14	7
5– 9	6
0– 4	0
Total	43

Mean..... 20.03 (27.4% of all the questions)
S.D..... 9.35
P.E..... 0.96

Table III shows the distribution of 92 scores on the Woodworth Questionnaire The mean score of this distribution is 17.83 unfavorable responses, which is 14.9 per cent of the 116 questions making up the test. The S.D. of the distribu-

tion is 9.75, P.E. 0.69. Thirteen individuals of this group score above the criterion of 30 unfavorable responses.

TABLE III
Distribution of scores on the Woodwarth Questionnaire

SCORE	FREQUENCY
40–44	2
35–39	6
30–34	5
25–29	6
20–24	13
15–19	16
10–14	23
5– 9	18
0– 4	3
Total	92

Mean	17.83
S.D.	9.75
P.E.	0.685

Comparison with other data:

The 116 questions of the original Woodworth Questionnaire can be divided into three groups; *i.e.* (1) those questions bearing directly upon the present condition of the individual; (2) indefinite present or historical; and (3) definite historical facts. Hollingworth[1] (3) labels these groups A, B and C respectively. Group A (present condition) contains 46 questions; group B (indefinite) contains 35 questions; and group C (historical) contains 35 questions. Comparative tables showing questions significantly answered by 25 per cent. or more of the cases will be discussed further on.

Hollingworth's distribution of scores for all cases including definitely diagnosed disorders of both pre- and post-armistice groups (psychoneuroses, hysteria, concussion, neurasthenia, constitutional psychopathy, and epilepsy) shows a

[1] Hollingworth, H. L. "The Psychology of Functional Neuroses," New York, 1920, p. 133.

median of 15–20 unfavorable responses. My group of 92 undiagnosed prison cases shows exactly the same median score; *i.e.* 15–20, or rather, 15–19.99 unfavorable responses The actual mean of Hollingworth's cases was 19.02 responses with a Standard Deviation of 15.60, P.E. 0.416.[2] The actual mean of my prison group was 17.83, P.E. 0.685, showing a slightly lower mean score. The median score of Hollingworth's pre-armistice group was 20–25 responses as contrasted with his post-armistice group which dropped to a median of 10–15 responses. However but 9.40 per cent of all the cases in the pre-armistice group occurred at the median as contrasted with the post-armistice group which included 26.24 per cent of the cases at the median for that class. For both groups combined 10.34 per cent of the cases occur at the median and for my prison group 17.39 per cent of the cases fall at that level. Hollingworth's combined group shows that 53.29 per cent stand above the median. My prison group shows that 47.83 per cent are below and 34.78 per cent are above the median. For the pre-armistice group 31.62 per cent fall below the median and 46.15 per cent are above. The post-armistice group shows 39.60 per cent below and 34.16 per cent above the median.

The range of scores in Hollingworth's combined groups is from 0–65 and over, whereas the prison group range is somewhat more restricted; *i.e.* 0–45. This is also evidenced by the difference in sigma values of the comparative groups. Table IV is a comparison of the two sets of data with various measures of variability calculated. The distribution of scores for the prison group seems to have the greatest amount of symmetry when plotted histographically although there is a slight amount of positive skewness. The combined pre-armi-

[2] The S.D., P.E., means, quartiles, and interquartile ranges of Hollingworth's data were not included in his study but were calculated from his tables. P.E.'s were calculated according to the formula:

$$\text{P.E.} = 0.6745 \frac{\sigma}{\sqrt{2N}}.$$

stice and post-armistice groups approach more of a J curve and is quite asymmetrical. The post-armistice group alone is of the same type. The pre-armistice group does not approach the characteristics of any of the other three. It is more elongated and plateau-like in character showing a more pronounced piling up of scores in the upper ranges of in-

TABLE IV

Showing comparison between the Prison Group and the Army Groups

SCORES	PRISON GROUP		ARMY GROUPS					
			PRE-ARMISTICE		POST-ARMISTICE		COMBINED GROUPS	
	No. of Cases	Per Cent	No. of Cases	Per Cent	No. of Cases	Per Cent	No. of Cases	Per Cent
65–	.....	.....	1	.85	1	.49	2	.62
60–65	.....	.....	1	.85	2	.99	3	.94
55–60	.....	.....	2	1.71	3	1.49	5	1.57
50–55	.....	.....	5	4.27	1	.49	6	1.88
45–50	.....	.....	6	5.13	12	5.94	18	5.64
40–45	2	2.17	10	8.55	3	1.49	13	4.08
35–40	6	6.52	7	5.98	2	.99	9	2.82
30–35	5	5.43	8	6.84	5	2.48	13	4.08
25–30	6	6.52	14	11.97	10	4.95	24	7.52
20–25	13	14.13	11	9.40	12	5.94	23	7.21
15–20	16	17.39	15	12.82	18	8.91	33	10.34
10–15	23	24.99	8	6.84	53	26.24	61	19.12
5–10	18	19.57	19	16.24	30	14.85	49	15.36
0– 5*	3	3.26	10	8.55	50	24.75	60	18.81
Totals	92	99.98	117	100.00	202	100.00	319	99.99
Q_3	23.46		36.96		20.21		27.76	
Mean	17.83		24.81		15.97		19.02	
Median	15.63		22.95		11.98		14.17	
Q_1	10.44		10.16		5.08		7.02	
S.D.	9.75		15.98		13.69		15.60	
P.E.	0.69		0.70		0.46		0.42	
Semi-interquartile range	6.51		13.40		7.47		10.37	

* The scores 0–5, 5–10, etc., mean 0–4.99, 5–9.99, etc.

stability. The prison group, then, is more compact and shows a slightly higher frequency of instability in the moderate ranges; *i.e.* 5–10, 10–15, 15–20 and 20–25, 76 per cent falling within that area of the curve. For the pre-armistice, post-armistice and combined groups, the percentage of cases within the same range; *i.e.* 5–20, is 45.3 per cent, 55.9 per cent and 52 per cent respectively.

Let us now look at the types of symptoms responsible for the rather high emotionality among the prison group. Table V shows the frequency of significant responses to questions of Class A (present condition) on the Cady Modification by 25 per cent or more of the cases. Refer to end of article for complete questionnaires and exact questions.

TABLE V

Frequency of response to Class A questions on Cady Modification by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
1	11	25.0
3	21	47.7
9	11	25.0
10	18	41.0
17	16	36.4
23	11	25.0
27	13	29.3
39	18	41.0
41	17	38.6
51	16	36.4
54	18	41.0
59	15	34.1
61	19	43.2
71	25	57.0
77	16	36.4
79	17	38.6
80	22	50.0
82	20	45.5

The largest percentage of unfavorable responses in this classification concern themselves with: becoming tired of

people quickly, being generally regarded as disobedient, pains in the eyes, easily angered, feelings of unreality, phantasy life, worry over condition of health, difficulty in making decisions, abnormal dreams, feelings of suicide and hopelessness, tendency to prefer the company of older people and active imaginative life.

Table VI brings to light the frequency of significant responses to questions belonging to Class C (historical facts) by 25 per cent or more of the men.

TABLE VI

Frequency of responses to Class C questions (historical facts) on the Cady Modification by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
4	32	72.7
7	25	57.0
14	26	59.1
22	13	29.6
34	19	43.2
37	20	45.5
40	44	100.0
46	14	31.8
49	23	52.2
72	11	25.0

The above responses show an early desire to leave school, nicknames they didn't like, running away from home, teachers continually reprimanding them, expulsion from school, unjust punishments, arrests, inability to get on with teachers, feelings of wickedness, finding school a difficult place to get along in.

Table VII shows frequency of significant responses to indefinite questions (Class B) by 25 per cent or more of the men.

These responses are associated with getting out of breath quickly, desire for praise, ideas of suspicion, desire to wear expensive jewelry, rings, etc., inability to eat certain foods, inability to work with others looking on, feelings of unpopu-

TABLE VII

Frequency of responses to Class B questions (indefinite) on the Cady Modification by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
29	16	36.4
43	20	45.5
50	11	25.0
52	11	25.0
74	27	61.4
76	17	38.6
84	14	31.8
85	13	29.3

larity and inferiority, and feeling that they are not getting a square deal in life.

Questions on the Woodworth Questionnaire are analyzed in the same way and are included in Tables VIII, IX and X. With this tabulation we can make comparisons with Hollingworth's data, although we must constantly keep in mind that such comparisons of two distinct groups of individuals as these will not necessarily bring to light any very common factors of emotionality, because the motivating factors underlying the atypical behavior of each group is not comparable.

TABLE VIII

Frequency of responses to Class A questions (present condition) on the Woodworth Questionnaire by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
8	26	27.6
10	38	40.0
59	33	35.0
68	25	26.6
84	27	28.8
92	27	28.8
93	26	27.6
94	25	26.6
95	31	33.0
97	50	53.0

The above questions pertain to sleep disturbances, worries, desire to steal things, quick and easy change of moods, becoming tired of people and things quickly, and frequent change of interests.

TABLE IX

Frequency of responses to Class C questions (historical facts) on the Woodworth Questionnaire by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
31	34	36.2
32	52	44.7
36	60	64.0
37	65	69.0

Adolescent unhappiness, being considered a bad boy, running away from home characterize the above types of questions responded to significantly.

TABLE X

Frequency of significant responses to Class B questions (indefinite) on the Woodworth Questionnaire by 25 per cent or more of the cases

NO. OF QUESTION	FREQUENCY OF RESPONSE	PER CENT
12	25	26.6
13	28	29.8
16	33	35.0
43	25	26.6
49	27	28.8
69	34	36.2
115	46	49.0

Positive responses to the above enumerated questions concern pains in the eyes, sensations of mistiness before the eyes, frequent blushing, failure to make friends easily, feeling that use of tobacco has had an injurious effect, worries about unfinished tasks, and aversion of disagreeable odors.

Hollingworth averages the percentage of responses for each diagnostic group in each of the classes of questions. Table XI below shows comparative percentages of the prison group

with all diagnostic groups of Hollingworth both before and after the armistice.

TABLE XI
Comparative Results

Questions in Class A refer to Present Condition of Individual.
" " " B refer to Indefinite Present or Historical.
" " " C refer to Definite Historical Facts.

Figures are average percentages of significant responses to each class

CLASS	PRISON GROUP	HOLLINGSWORTH'S ARMY CASES	
		Pre-Armistice Group	Post-Armistice Group
A	32.7	42.0	17.8
B	33.1	40.2	21.3
C	53.5	35.2	24.5
Cases	92	95	179

The above table shows that the prison group studied does not compare closely with Hollingworth's cases, but lies more or less between the Pre-Armistice and Post-Armistice percentages with the exception of responses to the C questions in which the prison group responds much more positively. This is probably due to the fact that the prison group responded 25 per cent or more to only one question (unhappiness between 14 and 18 years of age) included in the responses to this class of questions by Hollingworth's cases. The four questions, as enumerated above in Table IX, are all more or less associated with delinquency and not necessarily with general instability. The traits thus shown are adolescent unhappiness, having been considered a bad boy, and running away from home and desire to run away from home.

Hollingworth's group are characterized in this class of questions (historical matters of fact) by fits of dizziness, fainting, hysterical blindness or deafness, loss of memory, unhappy childhood and adolescence, mental shock, biting of finger nails, motor tics and habit spasms, and nervous breakdown. We can thus see that the emotionality of Holling-

worth's group is truly more representative of psycho-neurotic symptoms, *per se,* than the prison group studied, especially for the class C questions.

Let us now look at questions in the A class (present condition). Hollingworth's group of army men respond positively 25 per cent or more to 23 of the questions in this class. The prison group respond to only 10 of the questions. (See Table VIII). Sleep disturbances, worries, desire to steal, quick change of moods, frequent change of interests, boredness, etc., are characteristic of this group of men. On the other hand, the army group are more characterized by general feelings of illness, sleep disturbances, night terrors, dreams, tiredness, cardiac disturbances, bodily pains (undifferentiated), bad headaches, feelings of not being their old self, not knowing what to do next, worry over trifles, trouble in making up the mind, easily rattled, fidgeting, mind wandering, useless thoughts running continually through the mind, perseveration, and easy to become angered.

From the above characteristics we see very little overlapping of the two groups; only 5 questions answered significantly by 25 per cent or more by both groups, *viz.;* two pertaining to sleep disturbances, feelings of not being old self, worry over trifles, and ease in becoming angered. The questions responded to by the army group seem to be more definite and diagnostic.

Questions in Class B (Indefinite Present or Historical) characterize the prison group by pains in the eyes, frequent blushing, failure to make friends easily, feeling that tobacco has had injurious effects, worries about unfinished tasks and aversion to disagreeable odors. The army group respond with more overlapping in this class with the prison group. The questions answered significantly by both groups in this class pertain to ocular disturbances, feeling that tobacco has been injurious, worry over unfinished tasks, and inability to stand disgusting smells. Other questions peculiar to the army group comprise cardiac disturbances, feelings of suffocation,

pressure in or about the head, body pains, inability to adjust to new surroundings quickly, uneasiness in going through a tunnel or subway, inability to do good work while others are looking on, inability to stand pain quietly or stand the sight of blood.

Even in this class of questions which is more indefinite than the other two classes, there seems to be a more positiveness and definiteness of symptomatology. Of course the army group is three times the size of the prison group, except the Pre-Armistice group (the most diagnostically unstable) which is approximately the same size. (See Table XI.) With the addition of another larger group (Post-Armistice) the chances for enlarging the number of questions is naturally increased, but an analysis of Hollingworth's data shows that the Post-Armistice group does not apparently affect the number of questions responded to significantly by 25 per cent or more of the Pre-Armistice group.

In summary it might be definitely stated that the army group of men referred to the psychopathic hospital for observation exhibited a more pronounced and clear-cut symptomatology than the prison group which was chosen more or less at random. The symptoms of the prisoners diagnostic of instability were on the whole more vague and indefinite in character.

Can we then say that any arbitrary number of significant responses to the Woodworth Questionnaire or similar modification is diagnostic of instability? The opinion of the writer is that any such designation such as Franz (1) mentions by adopting an arbitrary number of significant responses; *i.e.*, 20 or 30 out of the 116 questions comprising the questionnaire, must be carefully analyzed as to the type of question responded to. Could an individual responding significantly to 30 more or less vague questions be considered unstable to the same degree as another individual responding to 30 more positive and definite questions? In making our analyses and

diagnoses it seems best to consider the qualitative aspect of the responses rather than the quantitative.

Tables XII and XIII below give a detailed analysis of both the Cady Modification and the original Woodworth Questionnaire showing frequency of response to each question with percentages. Questions marked with an asterisk (*) show those responded to significantly by 25 per cent or more of the men. Reference to the actual questionnaires can be had at

TABLE XII

Frequency of significant responses on the Cady Modification of the Woodworth Questionnaire

QUES. NO.	FREQUENCY	PER CENT	QUES. NO.	FREQUENCY	PER CENT	QUES. NO.	FREQUENCY	PER CENT
*1 ..	11	25.0	*29	16	36.4	*59	15	34.1
2 ..	3	6.8	31	5	11.3	60	3	6.8
*3 ..	21	47.7	32	4	9.1	*61	19	43.2
*4 ..	32	72.7	33	4	9.1	62	10	22.8
5 ..	2	4.5	*34	19	43.2	64	9	20.4
*7 ..	25	57.0	35	10	22.8	65	1	2.3
8 ..	3	6.8	*37	20	45.5	66	3	6.8
*9 ..	11	25.0	38	2	4.5	67	10	22.8
*10 ..	18	41.0	*39	18	41.0	69	6	13.6
11 ..	5	11.3	*40	44	100.0	70	3	6.8
13 ..	8	18.2	*41	17	38.6	*71	25	57.0
*14 ..	26	59.1	*43	20	45.5	*72	11	25.0
15 ..	8	18.2	44	7	15.9	73	10	22.8
16 ..	8	18.2	45	4	9.1	*74	27	61.4
*17 ..	16	36.4	*46	14	31.8	75	7	15.9
19 ..	7	15.9	47	9	20.4	*76	17	38.6
20 ..	10	22.8	*49	23	52.2	*77	16	36.4
21 ..	5	11.3	*50	11	25.0	78	4	9.1
*22 ..	13	29.6	*51	16	36.4	*79	17	38.6
*23 ..	11	25.0	*52	11	25.0	*80	22	50.0
25 ..	7	15.9	*54	18	41.0	81	10	22.8
26 ..	1	2.3	55	1	2.3	*82	20	45.5
*27 ..	13	29.3	56	10	22.8	83	9	20.4
28 ..	4	9.1	57	2	4.5	*84	14	31.8
						*85	13	29.3

* Questions answered significantly by 25 per cent or more of the cases.

TABLE XIII

Frequency of significant responses on the Woodworth Questionnaire

QUES. NO.	FREQUENCY	PER CENT	QUES. NO.	FREQUENCY	PER CENT	QUES. NO.	FREQUENCY	PER CENT
1	7	7.5	40	6	6.4	79	11	11.7
2	10	10.6	41	11	11.7	80	3	3.2
3	1	1.1	42	10	10.6	81	18	19.1
4	8	8.5	*43	25	26.6	82	6	6.4
5	6	6.4	44	19	20.2	83	1	1.1
6	13	13.8	45	17	18.2	*84	27	28.8
7	0	0.0	46	6	6.4	85	21	22.4
*8	26	27.6	47	7	7.5	86	4	4.3
9	16	17.0	48	10	10.6	87	6	6.4
*10	38	40.0	*49	27	28.8	88	13	13.8
11	14	14.9	50	15	16.0	89	20	21.3
*12	25	26.6	51	8	8.5	90	2	2.1
*13	28	29.8	52	3	3.2	91	15	16.0
14	7	7.5	53	10	10.6	*92	27	28.8
15	4	4.3	54	8	8.5	*93	26	27.6
*16	33	35.0	55	1	1.1	*94	25	26.6
17	12	12.8	56	9	9.6	*95	31	33.0
18	6	6.4	57	7	7.5	96	23	24.4
19	20	21.3	58	16	17.0	*97	50	53.0
20	11	11.7	*59	33	35.0	98	21	22.4
21	9	9.6	60	21	22.4	99	11	11.7
22	16	17.0	61	10	10.6	100	7	7.5
23	10	10.6	62	0	0.0	101	4	4.3
24	12	12.8	63	2	2.1	102	1	1.1
25	19	20.2	64	3	3.2	103	1	1.1
26	1	1.1	65	1	1.1	104	1	1.1
27	8	8.5	66	9	9.6	105	0	0.0
28	3	3.2	67	21	22.4	106	9	9.6
29	11	11.7	*68	25	26.6	107	4	4.3
30	20	21.3	*69	34	36.2	108	7	7.5
*31	34	36.2	70	20	21.3	109	6	6.4
*32	42	44.7	71	19	20.2	110	3	3.2
33	13	13.8	72	21	22.4	111	0	0.0
34	0	0.0	73	10	10.6	112	18	19.1
35	3	3.2	74	12	12.8	113	10	10.6
*36	60	64.0	75	15	16.0	114	9	9.6
*37	65	69.0	76	2	2.1	*115	46	49.0
38	19	20.2	77	15	16.0	116	1	1.1
39	6	6.4	78	10	10.6			

* Questions answered significantly by 25 per cent or more of the cases.

the end of the article. The first three appended references also describe these questionnaires. The Cady Modification can also be found in "Genetic Studies of Genius," Vol. I, by Lewis M. Terman. The original Woodworth Questionnaire is included both by Franz (1) and Hollingworth (3).

SUMMARY AND CONCLUSIONS

Any definite conclusions concerning the general emotionality of the prisoner must of necessity be deferred until many more cases are tested. Further studies of this nature should also include more representative age groups as well as other types of testing devices. However, certain tentative conclusions can be drawn from the data presented in this report.

1. The number of cases tested in the survey is probably sufficient for comparative studies with data previously collected by other investigators using the Woodworth Personal Data Sheet or Questionnaire.

2. The use of the Woodworth Questionnaire was satisfactory for the purposes used in the survey, *viz.*, the selection of a small group of unstable and neurotic men.

3. There seems to be no valid reason for the objections to the use of this questionnaire for such purposes of selection.

4. A modification of the Woodworth Questionnaire was first used, but later rejected on account of its juvenile nature.

5. The distribution of scores on the modified form showed a mean score which was out of proportion to the number of questions comprising the total questionnaire when compared with the mean score attained on the original Woodworth form.

6. The criterion of selection was set at 30 unfavorable responses for both forms in order to compensate for the more or less youthful nature of the modified form.

7. 135 cases were run in order to get a group of 20 who scored 30 or more unfavorable responses.

8. The distribution of scores of the prison group is more symmetrical and has less variability than the groups studied by Hollingworth.

9. In general, the comparative data show similar results, but analytical treatment shows a much wider variability.

10. The types of symptoms characteristic of the unfavorable responses by the prison group seem to be more vague and indefinite than the unfavorable responses given by Hollingworth's groups. The latter seem to show a more definite and clear-cut symptomatology diagnostic of emotional instability.

11. The army group was a selected unstable group referred for treatment, whereas the prison group was a more or less random sampling, consequently the symptomatology expressed by the two groups is not comparable as the motivating factors underlying the causes of the instability are not the same for both groups.

12. Analyses of the two sets of data are more similar from a quantitative point of view than a qualitative.

13. A purely arbitrary number of unfavorable responses used as a criterion for the selection of unstable groups is not good procedure unless considered in the light of their qualitative aspects.

REFERENCES

1. Franz, S. I. Handbook of Mental Examination Methods, New York, 1919, pp. 193.
2. Cady, V. M. The Estimation of Juvenile Incorrigibility, Jour. of Delinquency, Monograph No. 2, April, 1923.
3. Hollingworth, H. L. The Psychology of Functional Neuroses, New York, 1920, pp. 259.
4. Matthews, Ellen. A Study of Emotional Stability in Children, Jour. of Delinquency, 8, 1, Jan. 1923, p. 1–40.
5. Flemming, Edwin G., and Flemming, Cecile White. The Validity of the Matthews' Revision of the Woodworth Personal Data Questionnaire, Jour. of Abnormal and Social Psychology, 23, 4, Jan.–March, 1929, p. 500–506.
6. Flemming, C. W. A Detailed Analysis of Achievement in the High School, Teachers College, Columbia University Contributions to Education, No. 196, 1925.
7. Terman, L. M. A New Approach to the Study of Genius, Psychol. Review, 29, 4, p. 310.
8. Ball, R. J. An Objective Measure of Emotional Instability, Jour. of Applied Psychol., 13, 3, June, 1929, p. 226–256.

Woodworth-Cady Questionnaire

AGE: —— GRADE: ——— NAME: ————————————

Directions: Answer every question as truthfully and honestly as you can by drawing a line under the right answer, as shown in the samples.

SAMPLE 1: Are there seven days in a week? YES NO

The right answer is yes, so the word yes has a line under it.

SAMPLE 2: Do you sleep 15 hours a day? YES NO

Now draw a line under yes or no to tell whether you sleep 15 hours a day.

Below are other questions. Answer every one truthfully and honestly by drawing a line under yes or no. Do not skip any. Begin writing.

1. Are you happy most of the time? ..... YES NO
2. Are you afraid in the dark? ..... YES NO
3. Do you get tired of people easily? ..... YES NO
4. Were you anxious to get away from school and get a job? ..... YES NO
5. Have you often fainted away? ..... YES NO
6. Were you ever fond of playing with snakes? ..... YES NO
7. Did you ever have a nickname you didn't like very well? ..... YES NO
8. Do you usually enjoy your meals? ..... YES NO
9. Can you usually sit still without fidgeting? ..... YES NO
10. Do people say you are disobedient? ..... YES NO
11. Do you usually sleep well? ..... YES NO
12. Can you keep still when you are being tickled? ..... YES NO
13. Do you make friends easily? ..... YES NO
14. Did you ever run away from home? ..... YES NO
15. Have you the habit of biting your finger nails? ..... YES NO
16. Does the thought of hurting a person or animal give you pain? ..... YES NO
17. Do your eyes often pain you? ..... YES NO
18. Do you prefer to go without your breakfast? ..... YES NO
19. Would you rather play by yourself than with someone? ..... YES NO
20. Do you ever dream of robbers? ..... YES NO
21. Did you ever have the habit of twitching your head, neck, or shoulders? ..... YES NO
22. Did your teachers tell you that you are too noisy or talk too much? ..... YES NO
23. Do you hear easily when spoken to? ..... YES NO
24. Do you prefer rainy weather? ..... YES NO
25. Do you sometimes cry yourself to sleep? ..... YES NO
26. Do you mind crossing a bridge over water? ..... YES NO
27. Do you ever have the same dream over and over? ..... YES NO
28. Do you like to tease people till they cry? ..... YES NO
29. Do you get out of breath quickly when you run? ..... YES NO
30. Would you like to go to school on Saturdays? ..... YES NO
31. Did your teachers generally treat you right? ..... YES NO

32. Can you stand the sight of blood? ..... YES NO
33. Are you troubled with dreams about your play? ..... YES NO
34. Were you ever expelled from school or nearly expelled? ..... YES NO
35. Do you usually feel well and strong? ..... YES NO
36. Do you refuse candy when it is offered you? ..... YES NO
37. Have you often been punished unjustly? ..... YES NO
38. Do you have a great fear of fire? ..... YES NO
39. Do you often get cross over little things? ..... YES NO
40. Have you ever been arrested, fined, or placed on probation? ..... YES NO
41. Do things ever seem to get misty before your eyes? ..... YES NO
42. Do you like to have people hurry you when you are working? ..... YES NO
43. Do you like to be praised and made much of? ..... YES NO
44. Can you stand as much pain as others can? ..... YES NO
45. Do you giggle a good deal and find you cannot help it? YES NO
46. Did you always get on well with your teachers or principal? ..... YES NO
47. Do you have many bad headaches? ..... YES NO
48. Is it easy for you to get up as soon as you awaken in the morning? ..... YES NO
49. Did you ever feel that you were very wicked? ..... YES NO
50. Do you know anybody who is trying to do you harm or hurt you? ..... YES NO
51. Are you ever bothered by a feeling that things are not real? ..... YES NO
52. Would you like to wear expensive jewelry, rings, etc.? ..... YES NO
53. Have you always remembered to clean your teeth? ..... YES NO
54. Do you like a make-believe life, in addition to your real life? ..... YES NO
55. Does it make you uneasy to cross a wide street or open square? ..... YES NO
56. Do you feel sort of tired a good deal of the time? ..... YES NO
57. Do those you play with say you quarrel or fight too much? YES NO
58. Do you dislike to ride in automobiles? ..... YES NO
59. Do you ever feel that no one loves you? ..... YES NO
60. Do you think you have more fears than most people? ..... YES NO
61. Do you sometimes worry about your health? ..... YES NO
62. Do you enjoy being idle and lounging about? ..... YES NO
63. Does the odor of most flowers make you sick? ..... YES NO
64. Do you think that people are always looking at you or making remarks about you? ..... YES NO
65. Do you have to have a light in your room at night before you can go to sleep? ..... YES NO
66. Do your folks say that you walk in your sleep? ..... YES NO
67. Do people find fault with you much? ..... YES NO
68. Would you be afraid to ride on horseback? ..... YES NO
69. Did other children of your age usually let you play with them? ..... YES NO
70. Are you often frightened in the middle of the night? ..... YES NO
71. Do you often have a hard time making up your mind about things? ..... YES NO
72. Did you find school a hard place to get along in? ..... YES NO

73. Are your feelings often hurt so badly that you cry? ..... YES NO
74. Is there any kind of food that disgusts you so that you cannot eat it? ..... YES NO
75. Did you ever have the habit of stuttering or stammering? YES NO
76. Can you do good work while people are looking at you? YES NO
77. Do you ever dream of people being dead? ..... YES NO
78. Can you handle yourself well in walking and running? ..... YES NO
79. Do you sometimes wish you had never been born? ..... YES NO
80. Would you rather be with grown people than with those your own age? ..... YES NO
81. Do you have trouble in getting to sleep? ..... YES NO
82. Do you ever imagine stories to yourself so that you forget where you are? ..... YES NO
83. Do you feel like jumping off when you are on a high place? ..... YES NO
84. Do you think people like you as much as they do other people? ..... YES NO
85. Do you feel that you are getting a square deal in life? ..... YES NO

When you have finished answering ALL the above questions, please hand in your paper to the examiner.

Woodworth Questionnaire

Answer every question as truthfully and honestly as you can by *drawing a line under* the right answer, as shown in the sample: SAMPLE: Are there seven days in a week? YES NO. The correct answer is YES, so the word YES is *underlined*. Do likewise for *all* the questions below. Do not skip any. Begin writing.

1. Do you usually feel well and strong? ..... YES NO
2. Do you usually sleep well? ..... YES NO
3. Are you often frightened in the middle of the night? YES NO
4. Are you troubled with dreams about your work? ..... YES NO
5. Do you have nightmare? ..... YES NO
6. Do you have too many sexual dreams? ..... YES NO
7. Do you ever walk in your sleep? ..... YES NO
8. Do you ever have the sensation of falling when going to sleep? ..... YES NO
9. Does your heart ever thump in your ears so that you cannot sleep? ..... YES NO
10. Do ideas run through your head so that you cannot sleep? ..... YES NO
11. Do you feel well rested in the morning? ..... YES NO
12. Do your eyes often pain you? ..... YES NO
13. Do things ever seem to swim or get misty before your eyes? ..... YES NO
14. Do you often have the feeling of suffocating? ..... YES NO
15. Do you have continual itchings in the face? ..... YES NO
16. Are you bothered much by blushing? ..... YES NO
17. Are you bothered by fluttering of the heart? ..... YES NO
18. Do you feel tired most of the time? ..... YES NO
19. Have you ever had fits of dizziness? ..... YES NO
20. Do you have queer, unpleasant feelings in any part of the body? ..... YES NO

21. Do you ever feel an awful pressure in or about the head? YES NO
22. Do you often have bad pains in any part of the body? YES NO
    Where? .....
23. Do you have a great many bad headaches? ..... YES NO
24. Is your head apt to ache on one side? ..... YES NO
25. Have you *ever* fainted away? ..... YES NO
26. Have you *often* fainted away? ..... YES NO
27. Have you ever been blind, half-blind, deaf or dumb for a time? ..... YES NO
28. Have you ever had an arm or leg paralyzed? ..... YES NO
29. Have you ever lost your memory for a time? ..... YES NO
30. Did you have a happy childhood? ..... YES NO
31. Were you happy when 14 to 18 years old? ..... YES NO
32. Were you considered a bad boy? ..... YES NO
33. As a child did you like to play alone better than to play with other children? ..... YES NO
34. Did the other children let you play with them? ..... YES NO
35. Were you shy with other boys? ..... YES NO
36. Did you ever run away from home? ..... YES NO
37. Did you ever have a strong desire to run away from home? ..... YES NO
38. Has your family always treated you right? ..... YES NO
39. Did the teachers in school generally treat you right? YES NO
40. Have your employers generally treated you right? ..... YES NO
41. Do you know of anybody who is trying to do you harm? YES NO
42. Do people find fault with you more than you deserve? YES NO
43. Do you make friends easily? ..... YES NO
44. Did you ever make love to a girl? ..... YES NO
45. Do you get used to new places quickly? ..... YES NO
46. Do you find your way about easily? ..... YES NO
47. Does liquor make you quarrelsome? ..... YES NO
48. Do you think drinking has hurt you? ..... YES NO
48. Do you think tobacco has hurt you? ..... YES NO
50. Do you think you have hurt yourself by going too much with women? ..... YES NO
51. Have you hurt yourself by masturbation (self-abuse)? YES NO
52. Did you ever think you had lost your manhood? ..... YES NO
53. Have you ever had any great mental shock? ..... YES NO
54. Have you ever seen a vision? ..... YES NO
55. Did you ever have the habit of taking any form of "dope"? ..... YES NO
56. Do you have trouble in walking in the dark? ..... YES NO
57. Have you ever felt as if some one was hypnotizing you and making you act against your will? ..... YES NO
58. Are you ever bothered by the feeling that people are reading your thoughts? ..... YES NO
59. Do you ever have a queer feeling as if you were not your old self? ..... YES NO
60. Are you ever bothered by a feeling that things are not real? ..... YES NO
61. Are you troubled with the idea that people are watching you on the street? ..... YES NO

62. Are you troubled with the fear of being crushed in a crowd? ..... YES NO
63. Does it make you uneasy to cross a bridge over a river? YES NO
64. Does it make you uneasy to go into a tunnel or subway? YES NO
65. Does it make you uneasy to have to cross a wide street or open square? ..... YES NO
66. Does it make you uneasy to sit in a small room with the door shut? ..... YES NO
67. Do you usually know just what you want to do next? YES NO
68. Do you worry too much about little things? ..... YES NO
69. Do you think you worry too much when you have an unfinished job on your hands? ..... YES NO
70. Do you think you have too much trouble in making up your mind? ..... YES NO
71. Can you do good work while people are looking on? ..... YES NO
72. Do you get rattled easily? ..... YES NO
73. Can you sit still without fidgeting? ..... YES NO
74. Does your mind wander badly so that you lose track of what you are doing? ..... YES NO
75. Does some particular useless thought keep coming into your mind to bother you? ..... YES NO
76. Can you do the little chores of the day without worrying over them? ..... YES NO
77. Do you feel you must do a thing over several times before you can drop it? ..... YES NO
78. Are you afraid of responsibility? ..... YES NO
79. Do you feel like jumping off when you are on a high place? ..... YES NO
80. At night are you troubled with the idea that somebody is following you? ..... YES NO
81. Do you find it difficult to pass urine in the presence of others? ..... YES NO
82. Do you have a great fear of fire? ..... YES NO
83. Do you ever feel a strong desire to go and set fire to something? ..... YES NO
84. Do you ever feel a strong desire to steal things? ..... YES NO
85. Did you ever have the habit of biting your finger nails? YES NO
86. Did you ever have the habit of stuttering? ..... YES NO
87. Did you ever have the habit of twitching your face, neck or shoulders? ..... YES NO
88. Did you ever have the habit of wetting the bed? ..... YES NO
89. Are you troubled with shyness? ..... YES NO
90. Have you a good appetite? ..... YES NO
91. Is it easy to make you laugh? ..... YES NO
92. Is it easy to get you angry? ..... YES NO
93. Is it easy to get you cross or grouchy? ..... YES NO
94. Do you get tired of people quickly? ..... YES NO
95. Do you get tired of amusements quickly? ..... YES NO
96. Do you get tired of work quickly? ..... YES NO
97. Do your interests change frequently? ..... YES NO
98. Do your feelings keep changing from happy to sad and from sad to happy without any reason? ..... YES NO
99. Do you feel sad or low-spirited most of the time? ..... YES NO

100. Did you ever have a strong desire to commit suicide? YES NO
101. Did you ever have heart disease? YES NO
102. Did you ever have St. Vitus's dance? YES NO
103. Did you ever have convulsions? YES NO
104. Did you ever have anemia badly? YES NO
105. Did you ever have dyspepsia? YES NO
106. Did you ever have asthma or hay fever? YES NO
107. Did you ever have a nervous breakdown? YES NO
108. Have you ever been afraid of going insane? YES NO
109. Have any of your family been insane, epileptic or feeble-minded? YES NO
110. Has any of your family committed suicide? YES NO
111. Has any of your family had a drug habit? YES NO
112. Has any of your family been a drunkard? YES NO
113. Can you stand pain quietly? YES NO
114. Can you stand the sight of blood? YES NO
115. Can you stand disgusting smells? YES NO
116. Do you like out-door life? YES NO

REMARKS: (Write on the reverse of this sheet anything you would like to say on any of the questions).

DETERMINING THE POINT OF EMPHASIS IN COLLEGE–ADULT LEVEL TESTING

R. A. BROTEMARKLE
University of Pennsylvania

The establishing of a greater reliability in our testing materials should come through a reasonably sympathetic consideration of the objections made by critics of testing procedure. The first step toward this is the locating of the point of emphasis in testing at which reasonable reliability is indicated.

Fernberger[1] has been preeminently fair in his conclusions about the necessary variability in mental tests. "Mental tests, as they are now developed, show such a degree of variability within a relatively homogeneous group that the differences between two groups do not have statistical significance, even though the differences may be great." "Such a degree of variability is the thing desired in a mental test—with the reduction of the variability one has a correspondingly great reduction in the value of the test for the purposes of differentiating the members of the group." The clinical psychologist is quite satisfied with his statement that, "this emphasizes the point-of-view that less weight is to be put on final test scores. From this viewpoint, mental tests become merely a standardized means of having the subject do something so that the trained examiner may observe his behavior and thus may arrive at a qualitative analytic diagnosis of the individual case."

But all of this drives the clinician into the necessity of establishing a higher degree of reliability for his diagnosis.

A study of five years' accumulation of results in the files of

[1] Fernberger, S. W., Statistical and Non-Statistical Interpretation of Test Results. Psychol. Clinic, 14, 3–4, pp. 2–5.

the College Personnel Office enables us to point toward the mental testing materials in which this reliability is most definitely indicated. Five forms of "mental" testing materials have been employed during this period. The Group tests have been represented by the Scholastic Aptitude Test of the *College Entrance Examination Board;* specific "mental ability" tests have been represented in the forms presented by Brotemarkle[2] in the presentation of analyses and standards. The analysis of these tests has established a division into certain Fundamental Ability measures of the biological functioning of the mental response, and Complex Mental Process measures of the intellectual organization of the mental function. The functional or achievement tests have been represented by certain measures of informational content, the use of amassed experience in relation to life experiences, motor control and emotional response.

The entire purpose of education at the college level in the development of personality is accomplished only upon the constant successful continuance of the student in the academic course set by the curriculum. Those who are unable to carry on momentarily but who may succeed in the end, faculties are prone to "place on general probation" and to limit their rosters until they have the opportunity to get going. Those who are entirely unable to accomplish the work are "dropped from the rolls" and eliminated at once from the educational process.

The testing ground of reliability for college-level mental tests lies in the distinguishing of the individuals in these groups in order to save the monetary cost of "educational waste," either by refusal of admission in the selective process, more rapid adjustment, or earlier elimination.

The following tables indicate the maximum and minimum limits and quartile scatter of these three groups of individ-

[2] Brotemarkle, R. A., College Student Personnel Problems, I. Individual Mental Testing at the College-Adult Level, II. The Analytical Study of the Student Personnel Problem, Jour. Appl. Psych., 11, 6, 1927, pp. 415–436, and 12, 1, 1928, pp. 1–42.

uals, collected from our files and studies over a period of years. The three groups are marked "Normal Progress," "General Probation" and "Dropped" to indicate the individuals referred to above. The scores used are those previously reported by Brotemarkle.[2]

TABLE 1
Memory Span for Digits (Fundamental Ability Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	12	12	12
Q_3 (75 percentile)	9	8	8
Median	8	7	7
Q_1 (25 percentile)	6	6	6
Minimum	4	5	5

TABLE 2
Memory Span for Sentences (Syllables) (Fundamental Ability Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	50	40	40
Q_3	30	30	30
Median	30	25	25
Q_1	30	25	25
Minimum	15	10	10

TABLE 3
Taylor Number Test (Attention) (Fundamental Ability Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	28	24	24
Q_3	16	16	16
Median	13	13	13
Q_1	11	11	11
Minimum	5	4	3

TABLE 4

Healy Completion II (Apperception) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	100	100	100
Q_3	73	75	75
Median	63	64	64
Q_1	52	53	53
Minimum	21	30	21

TABLE 5

Ferguson Formboard No. 4 (Intelligence) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	27	30	35
Q_3	61	60	59
Median	84	84	83
Q_1	120	120	120
Minimum	300 (F)	300 (F)	300 (F)

TABLE 6

Ferguson Formboard No. 5 (Intelligence) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	40	40	40
Q_3	84	77	81
Median	116	103	113
Q_1	170	155	161
Minimum	300 (F)	300 (F)	300 (F)

TABLE 7
Ferguson Formboard No. 6 (Intelligence) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	42	41	53
Q_3	124	118	128
Median	196	192	190
Q_1	283	287	300 (F)
Minimum	300 (F)	300 (F)	300 (F)

TABLE 8
Trabue Language Scale J (Languageability) (Complex Mental Process Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	14	13	14
Q_3	11	10	10
Median	10	10	9
Q_1	9	8	8
Minimum	4	4	2

TABLE 9
Roback Opposites Test (Discrimination) (Complex Mental Process Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	88	82	88
Q_3	58	48	48
Median	50	40	40
Q_1	42	32	28
Minimum	16	16	4

TABLE 10
Roback Judgment Test (*Judgment*) (*Complex Mental Process Series*)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	88	93	93
Q_3	68	55	54
Median	60	50	46
Q_1	50	42	37
Minimum	25	20	20

TABLE 11
Adult Scaled Information Test (*Information*) (*Functional Test Series*)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	19	20	20
Q_3	15	14	13
Median	13	12	11
Q_1	11	10	9
Minimum	4	3	1

TABLE 12
Roback Application Test (*Application*) (*Functional Test Series*)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	93	80	80
Q_3	52	43	36
Median	40	33	30
Q_1	30	23	20
Minimum	0	0	0

TABLE 13
Brotemarkle Comparison Test (Emotional Response) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	6	8	10
Q_3	20	22	24
Median	24	26	28
Q_1	30	34	36
Minimum	94	106	136

TABLE 14
Healy Motor Tapping Test (Motor Control) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	142	130	134
Q_3	113	106	105
Median	103	98	96
Q_1	92	88	87
Minimum	42	45	10

TABLE 15
Scholastic Aptitude Test (Educability) (Functional Test Series)

	NORMAL PROGRESS	GENERAL PROBATION	DROPPED
Maximum	100	100	100
Q_3	70	52	50
Median	47	29	24
Q_1	23	12	8
Minimum	1	1	1

The variability within these groups, as Fernberger[1] indicates, makes the statistical study of results meaningless. But certainly some of the tests indicate reasonably high reliability in the differentiation of the groups. When inter-quartile changes are observed there is apparently a definite reliability of measure. This change is not observed among the Fundamental Ability Series Tests—which measure largely the mental activity of the individual on the biological level of mental response—except in highly integrated materials such as the Memory Span for Sentences. The continued use of such tests at the college-adult level is indicated only for the complete understanding of the picture of mental function. This has been detailed previously by Brotemarkle.[2]

This change is not observed in the simpler Functional Test Series, as indicated by the Healy Completion II in the study of the apperceptive process; nor in the more difficult Functional Test Series, as indicated by the Ferguson Formboards in the testing of intelligent solution of the problem at hand. These tests must be reserved for the same use as the Fundamental Ability Series tests indicated above.

However, such a definite change of standing in ability is indicated in the Complex Mental Processes which tap the intellectual function of the organized mental processes; and the more complex Functional Series tests which tap the intellectual adjustment of the individual to complex situations. Language ability, discrimination, and judgment clearly reveal such a difference between the groups; while information measures and application of amassed information are still more indicative. When a battery test such as the Scholastic Aptitude Test is employed the differences of complex mental function are even more pronounced. To this must be added the weight of the measures of emotional stability in individual and social adjustment of personality. The uncertain Comparison Test of the Ideational Content of Moral Concepts clearly differentiates between these groups. In light of the small standard deviation within the tests of this battery, the

differences are pronounced. This latter picture is emphasized by the instability qualitatively observed in the Motor Control Test. No other explanation can account for the variations within this single test device.

On the whole, the point of emphasis in mental testing at the college-adult level is clearly indicated as ranging through the more complex mental processes and their functional use in stabilized emotional response. This necessitates the continued use of such personnel procedures as will coordinate all of the personal behavior of the individual on the problem of academic success. Whether such evidences be gathered by the Student Personnel Questionnaire as reported by Brotemarkle[2] and subsequent interviews or by the direct interview methods is not so important; the coordinated use of the facts is the essential necessity for the proper solution of student personnel problems.

CONCLUSIONS

(1) The desired variability in mental testing at the college-adult level removes the possibility of statistical treatment usually demanded of experimental results.

(2) The reliability of these materials must be established in light of the requirements made on them by the function of college education.

(3) The point of emphasis in mental testing at the college-adult level is clearly indicated as lying among the complex mental processes of the intellectual organization of mental function, the functional use of these processes in complex problems, and the measures of stability in the emotional reactions of individual and social adjustment.

STUDIES OF TYPOGRAPHICAL FACTORS INFLUENCING SPEED OF READING

VII. VARIATIONS IN COLOR OF PRINT AND BACKGROUND[1]

MILES A. TINKER AND DONALD G. PATERSON
University of Minnesota

The increased use of color in magazine advertising, posters, car cards, booklets, and circulars, especially during the past ten or fifteen years, emphasizes the importance of scientific knowledge concerning the effect of various color combinations of type and background on speed of reading.[2] Colored print which is legible on one background may prove to be illegible when combined with certain other colored backgrounds.

Although recent texts dealing with the psychology of advertising devote a large amount of space to the topic of color, most of them treat attention value, color preference, and appropriateness of color combinations for particular purposes, and mention legibility only briefly. Presumably neglect of the legibility factor is due to the absence of experimental data. The only results cited are those briefly reported by Luckiesh for printed matter read from a distance.[3] He lists thirteen

[1] The writers gratefully acknowledge their indebtedness for aid in conducting these studies to the following: J. B. Lippincott Co., Philadelphia, for permission to reprint the Chapman-Cook Speed of Reading Tests, Forms A and B, for experimental purposes; the University of Minnesota Graduate School for research grant; and the psychology departments of Ohio State University and Washington University for permission to test psychology students.

[2] C. R. Thompson, The Use of Color in Magazine Advertisements, M. A. Thesis, Graduate School, University of Minnesota, July, 1930; also D. Starch, Principles of Advertising, A. W. Shaw Co., Chicago, 1923, pp. 579–583.

[3] M. Luckiesh, Light and Color in Advertising and Merchandising, H. Van Nostrand Co., N. Y., 1923, pp. 246–251.

combinations of colored print and background in order from most legible to least legible. Detailed statements regarding the conditions of the experiment such as number of subjects, kind of ink and paper employed, size of type, line width, text used, etc., are omitted. Furthermore, the rank order as given affords no indication of the amount of differences in legibility between the several combinations. For these reasons more definite information is needed as a guide to improved advertising practice.

METHOD

Although a very large number of color combinations might have been utilized in an experiment of this kind it was necessary to limit the present study to a small number of combinations. The four most legible and the three least legible combinations were selected from the list reported by Luckiesh. Inspection of available colored paper stocks led to the choice of three additional color combinations. Since black on white was used as a standard, data for a total of eleven color combinations are available.

The Chapman-Cook Speed of Reading tests were utilized in the identical manner described in previous articles in this series.[4] Form A was printed with black ink on white Rainbow cover-stock and served as the standard in each comparison. Form B was printed with Ruxton's colored ink on Rainbow cover-stock in the variations shown in Table I.

The black on white is really black on light grayish white cover-stock. As nearly as we could tell by comparing, under daylight illumination, the various test blanks with the Milton Bradley Company's Studio Book of Colored Papers (8165: 11–1–25), the colors of the various cover-stocks approach maximum saturation (chroma). In column 3 of Table I the attempt has been made to indicate brightness (value) as well as hue (color).

[4] See "Studies of Typographical Factors Influencing Speed of Reading" which have appeared in previous issues of J. Appl. Psychol.

TABLE I

Test Groups, Color Combinations of Ink and Paper, and Observed Color Effects

Column 1 shows test group, column 2 trade names for Ruxton's ink and Rainbow cover-stock used, and column 3 the observed effects of colored ink and paper combinations

Test Group	Trade Names	Observed Effect
(1)	(2)	(3)
I.	Grass green on white	Dark green on light grayish white
II.	Lustre blue on white	Dark blue on light grayish white
III.	Black jobbing on yellow	Black on yellow (slight orange tinge)
IV.	Tulip red on yellow	Light red on yellow (slight orange tinge)
V.	Tulip red on white	Light red on light grayish white
VI.	Grass green on red	Dark grayish green on red (dark tint)
VII.	Chromium orange on black	Dark lemon yellow on dark grayish black
VIII.	Chromium orange on white	Light orange on light grayish white
IX.	Tulip red on green	Dark brown on dark green
X.	Black jobbing on purple	Black on dark purple (violet)

All test blanks were printed in Scotch Roman, 10-point type, 19 pica line width, set solid. Data are reported for 85 college students in each group, or a total of 850 subjects in all. Of these, 288 were elementary and advanced psychology students at Washington University and 562 were elementary psychology students at Ohio State University.

RESULTS

Table II presents the detailed data for each comparison. Results for the ten test groups have been arranged in order from the color combination showing the least difference from black on white, to the one showing the greatest difference.

TABLE II

Influence of Different Combinations of Colored Print and Colored Paper

Differences given are the mean score on Form A using black print on white Rainbow cover-stock, minus the mean score on Form B varying with respect to color of ink and Rainbow cover-stock used. All test blanks were printed in Scotch Roman, 10 point type, 19 pica line width, set solid. The mean score is the average number of paragraphs of 30 words each in the Chapman-Cook Speed of Reading Tests read in 1¾ minutes. In each sub-group, N = 85 college students.

Test Group	Test Form, Color of Ink and Cover-stock	Mean	P.E.$_{dist.}$	P.E.$_{M}$	DIFFERENCE BETWEEN MEANS IN			P.E.$_{diff.}$	$\frac{D}{P.E._{diff.}}$
					Paragraphs	Words per Sec.	Per Cent		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
I	A, Black on White	17.26	2.94	.32	0.51	0.15	3.0	.16	3.24
	B, Green on White	16.75	2.58	.29					
II	A, Black on White	16.61	2.65	.29	0.55	0.16	3.4	.17	3.24
	B, Blue on White	16.06	2.39	.26					
III	A, Black on White	16.37	2.90	.31	0.60	0.17	3.8	.15	3.89
	B, Black on Yellow	15.77	2.42	.26					
IV	A, Black on White	17.09	2.47	.27	0.79	0.23	4.8	.17	4.74
	B, Red on Yellow	16.31	2.25	.24					
V	A, Black on White	17.65	3.41	.37	1.45	0.41	8.9	.20	7.19
	B, Red on White	16.20	2.80	.30					
VI	A, Black on White	17.91	2.70	.29	1.71	0.49	10.6	.18	9.69
	B, Green on Red	16.20	2.39	.26					

TABLE II

Test Group	Test Form, Color of Ink and Cover-stock	Mean	P.E.$_{dist.}$	P.E.$_{M}$	DIFFERENCE BETWEEN MEANS IN			P.E.*$_{diff.}$	$\frac{D}{P.E._{diff.}}$
					Paragraphs	Words per Sec.	Per Cent		
VII	A, Black on White	16.94	2.78	.30	2.01	0.57	13.5	.17	12.05
	B, Orange onBlack	14.93	2.15	.23					
VIII	A, Black on White	16.62	2.60	.28	2.87	0.82	20.9	.19	15.46
	B, Orange on White	13.75	2.53	.27					
IX	A, Black on White	17.38	2.50	.27	4.92	1.41	39.5	.25	19.91
	B, Red on Green	12.46	2.82	.31					
X	A, Black on White	16.66	2.52	.27	5.66	1.62	51.5	.28	20.29
	B, Black on Purple	11.00	3.08	.33					

* P.E.$_{diff.}$ values in column 9 were computed on the basis of the standard formula for determining the statistical significance of obtained differences for *correlated measures*. The correlations for the ten sub-groups in this experiment were $+0.85 \pm .02$ for Group I; $+0.81 \pm .03$ for Group II; $+0.87 \pm .02$ for Group III; $+0.79 \pm .03$ for Group IV; $+0.84 \pm .02$ for Group V; $+0.80 \pm .03$ for Group VI; $+0.83 \pm .02$ for Group VII; $+0.78 \pm .03$ for Group VIII; $+0.64 \pm .04$ for Group IX; and $+0.59 \pm .05$ for Group X.

In every comparison black on white yields the fastest reading rate. The differences shown for test groups I to III inclusive are not statistically certain. In other words the direction of the differences found might be reversed occasionally (once or twice in a hundred times) in further repetitions of the experiment. Furthermore, these differences are slight, amounting to a differential advantage of less than four per cent in favor of black on white (ranging from 3.0 to 3.8 per cent). From a practical viewpoint one might say that green on white, blue on white, and black on yellow are almost as effective as black on white in producing optimal reading conditions.

From Group IV through Group X the differences show a steady increase in statistical significance. Red on yellow, Group IV, shows a 4.8 per cent differential, which is sufficient to suggest the disadvantage of such a color combination whenever reading speed is a factor to be considered. In the remaining groups, the various color combinations produce very illegible text. The last four color combinations produce such unfavorable reading conditions that it is inadvisable ever to use them.

The black on purple combination was originally selected to obtain a quantitative measure of the difference when readers are confronted with an extremely illegible color combination. It is somewhat surprising to find that an average score of 11.00 was obtained because many of the subjects gave up in their attempts to read the copy. As a matter of fact practically all of the students exhibited varying degrees of consternation when they turned over this test blank and attempted to read it as quickly and accurately as possible. The presence of these varying attitudes of dismay, amounting in some instances to strong emotional reactions, is reflected in the correlation of only +.59 between Form A and Form B for this test group.

DISCUSSION OF RESULTS

A comparison of the results obtained in this experiment with those reported by Luckiesh is shown in the following

summary (ranking only those combinations which were utilized in both studies):

Color Combinations	RANK ORDER ACCORDING TO:	
	Present Study	Luckiesh
Black on white	1	5
Green on white	2	2
Blue on white	3	4
Black on yellow	4	1
Red on yellow	5	6
Red on white	6	3
Green on red	7	7
Red on green	8	8

Aside from the discrepancy in ranks for black on white, black on yellow, and red on white, the two studies are in close agreement. This amount of agreement is surprising since we have little assurance that our color combinations are identical with those used by Luckiesh even though the same color names are reported. Color terminology is notoriously unstandardized and unsatisfactory. It is important to point out that in spite of unsatisfactory color terminology it is possible to perform experiments which will yield surprisingly consistent results.[5] Thus it is feasible to set up tentative standards which will be useful in selecting legible color combinations.

According to Poffenberger the explanation of the results reported by Luckiesh does not depend on *color* as such but on *brightness difference.*[6] He states: ". . . legibility depends upon relation of color to background and the all important factor is *brightness difference*" (p. 263). Inspection of the

[5] In view of the unsatisfactory terminology used by the printing and paper industry in assigning names to colors all studies involving colored ink and paper should report the trade names together with description of the color effects observed. Even so, such practice must be recognized as being a makeshift pending the development and general acceptance of a valid system of color notation.

[6] A. T. Poffenberger, Psychology in Advertising, A. W. Shaw Co., Chicago, 1925, pp. 262–265.

printed texts employed in the present experiment supports the above conclusion. For example, as shown in column 3 of Table I, the color combinations at the top of the table have a maximum brightness contrast, those in the middle section a medium brightness contrast, and those at the bottom have a minimum brightness contrast. This order corresponds directly to the speed of reading rates found.

How misleading color combinations may be, judged solely from knowledge of color values singly, is shown by the effects observed in this experiment. For example, what the printer designates as orange on black produces the appearance of dark lemon yellow on dark grayish black.[7] This same ink when applied to white background appears as light orange. The reader will note other similar instances in Table I.

SUMMARY

1. Absence of quantitative studies dealing with the effect of various color combinations on speed of reading led to the present study.

2. Ten color combinations in addition to black on white were compared with respect to their effect on reading rate. Ten groups of 85 college students each were tested—a total of 850 subjects.

3. Striking differences in rate of reading text material printed with various combinations of colored ink and paper were disclosed. The following list derived from the detailed tables is offered as a rough guide for the advertiser:

Providing good legibility: Black on white, grass green on white, lustre blue on white, and black on yellow.

[7] Results obtained in the present study hold only for colored inks printed on *colored backgrounds*. Printing processes in which both letters and background are printed with different colored ink so as not to overlap would produce different effects. For example, orange letters printed on white paper with a black background printed on the same white paper (two color printing) would undoubtedly present a marked brightness contrast, the orange colored ink appearing as orange rather than as lemon yellow.

Providing fair legibility: Tulip red on yellow, tulip red on white.

Providing poor legibility: Grass green on red, chromium orange on black, chromium orange on white, tulip red on green, black on purple.

4. The evidence in this experiment justifies the following rule: In combining colors (color of ink and paper) care must be taken to produce a *printed page* which shows a maximum *brightness contrast* between print and background.

CONCERNING THE MAGNITUDE OF THE BEHAVIOR SAMPLE FOR THE STUDY OF BEHAVIOR TRAITS IN CHILDREN

MAPHEUS SMITH
University of Kansas

The literature on the study of personality by means of observations has given little serious attention to the problem of the length of observation record needed for obtaining statistically reliable data for the study of individual differences. Students have either failed to notice the problem or have passed it by with brief mention. For example, Goodenough, although presenting an argument in support of the reliability of her methods, only obtained an average reliability coefficient of +.600.[1] While this may be sufficient when the range of traits is limited, it is hardly adequate for intensive studies in behavior consistency.

A more adequate statement has been made by Loomis, who says at least two hours and possibly more time may be needed to get satisfactory samples of types of activity. The exact magnitude of the samples must be determined by increasing the length of the observation record "until the observations when correlated with each other are fairly stable and reliable."[2]

Olson[3] has presented the only actual data toward a solution of this problem. He experimented with various numbers and

[1] Goodenough, F. L., "Measuring Behavior Traits by Means of Repeated Short Samples," *Journal of Juvenile Research*, Vol. 12 (1928), pp. 230–235.

[2] Loomis, A. M., "A Preliminary Study of the Physical Contacts of Nursery School Children," in Thomas, D. S. (Ed.), *Some New Techniques for Studying Social Behavior.* New York: Teachers College, Columbia University, Bureau of Publications, 1929, pp. 73–75.

[3] Olson, W. C., "The Incidence of Nervous Habits in Children," *Jour. of Abnorm. and Social Psychol.*, Vol. 35 (1930–31), pp. 75–92.

lengths of observations. For seven ten-minute periods the coefficient of correlation was .75, which would become a reliability coefficient of .86 for fourteen observations by use of the prediction formula. Fourteen ten-minute observations yielded a reliability coefficient of .93, but the conditions were more constant. Coefficients of reliability for ten ten-minute observations of .69 to .76 were obtained for children, while twenty five-minute observations produced an average reliability of .87 $\pm$ .01. Olson also stated that an increase in the number of observations will increase the reliability as desired.

In order to develop this problem somewhat further in the direction of statistical treatment, two boys, each between six and seven years of age, were observed for a total of 500 minutes each in periods of approximately five minutes' duration, the data being written in the simplest possible language, without the use of categories during the period of recording. The observations were made in the playroom and on the playground of a child placing institution during regular play periods.

The material on each child was analyzed and classified into forty-two categories in such a manner as to have mutually exclusive groupings of fifty minutes each for intercorrelations. The method of analysis was to count the occurrence of certain types of behavior such as expressions denoting pleasure, displeasure or indefinite feeling, facial movements, body gestures, looking at others, contact with others of the nature of striking and pushing, other contacts, mouth behavior, body movements, running, walking, riding, other locomotion, manipulation, types of speech such as commands, requests, questions, declarations, uses of "I" and "me," and range of interaction.[4] Some other categories were more comprehensive, such as amount of behavior occurring when the child was with adults, with children, with others and material ob-

[4] This concept was used to mean the number of other different individuals appearing in the record of the child being studied.

jects, with others without material objects, alone with material objects, and alone without material objects.[5]

Besides the intercorrelations for each child of each fifty minute group with every other such group, data for each of the five mutually exclusive 100 minute groupings were intercorrelated, as were the first three mutually exclusive 150 minute groupings, the first two mutually exclusive 200 minute groupings, and the two mutually exclusive 250 minute groupings. Spearman's rank difference formula was used for all correlations.

RESULTS

It is seen from Table I that the mean of the forty-five intercorrelations for groupings of fifty minutes, for case No. 1, was .83 with a mean probable error of $\pm$ .0311. The standard error of the mean of these intercorrelations was $\pm$ .012, and the standard error of the mean P.E. of the intercorrelations was $\pm$ .0016. This means that there are approximately 68 chances in 100 that the difference between .83 and the true mean of the coefficients of correlation of the fifty minute groupings on case No. 1 will not be greater than .012. Correspondingly, there are 68 chances in 100 that the difference between the mean P.E. of the coefficients of correlation of $\pm$ .0311 and the true mean of the probable errors will not be greater than .002. Furthermore, the true average will almost certainly fall within three times the standard errors of .83 and $\pm$ .0311, respectively, or between .794 and .866 for the mean of the coefficients of correlation, and between .02511 and .03711 for the mean of the probable errors of the coefficients of correlation. Similar significance is attached to the means, probable errors, and standard errors of the other groupings for case No. 1 and to the means, probable errors, and standard errors of the groupings for case No. 2.

[5] The last four categories are after Arrington in Thomas, D. S. (Ed.) *Some New Techniques for Studying Social Behavior*, mentioned above, pp. 52–54.

TABLE I

NO. MIN. IN GROUP	CASE NO. 1				CASE NO. 2			
	N	MEAN r MEAN P.E.	σ AVE. MEAN	σ AVE. P.E.	N	MEAN r MEAN P.E.	σ AVE. MEAN	σ AVE. P.E.
50	45	.83 ± .0311	.012 ± .0016		45	.86 ± .0331	.0077 ± .0012	
100	10	.88 ± .026	.0156 ± .0027		10	.89 ± .022	.0080 ± .0019	
150[6]	3	.92 ± .017			3	.913 ± .017		
200[6]	1	.94 ± .01			1	.92 ± .01		
250[6]	1	.92 ± .01			1	.94 ± .01		

N is equivalent to number of intercorrelations.

Mean r is equivalent to mean coefficient of correlation.

Mean P.E. is equivalent to mean of the probable errors of the coefficients of correlation.

σ ave. Mean is equivalent to standard error of the mean.

σ ave. P.E. is equivalent to standard error of the mean P.E.

[6] The frequencies here are too small for significant treatment in terms of standard errors.

Although lack of numbers of groupings of 150, 200, and 250 minutes' duration made it useless to compute standard errors for those groupings, the increase of the mean correlation with the increase on the magnitude of the sample is clear. The use of the prediction formula of Spearman—$r_x = \frac{Nr}{1 + (N\text{-}1)r}$—in connection with the data indicate that, for case No. 1 to attain a reliability of .95 the length of the sample must be about 185 minutes, that for a reliability of .99 the sample must be 1,015 minutes, while it is impossible ever to get the perfect correlation of 1.00, which could insure that within a group of cases each trait of each child would be such that its rank in the group would probably not be changed.

Case No. 2 would need somewhat less observation. Since it possesses a higher mean r for fifty minutes, a reliability of .95 should be reached at 155 minutes and a reliability of .99 at 800 minutes. This case, in short, is that of a child whose behavior is more consistent in similar situations than is true of case No. 1.

Comparison of the actual mean r for 100 minutes and 150 minutes for these two cases with the values obtained from the prediction formula show that the mean r does not increase as predicted. The data are as follows: Case 1: Mean r at 50 minutes .83, at 100 minutes .88, at 150 minutes .92, where the prediction formula shows .83, .91, and .94 for similar groupings; Case 2: Mean r at 50 minutes equals .86, at 100 minutes .89, and at 150 minutes .913, where the prediction formula shows .86, .925, and .949. The presence of other factors, such as change in behavior due to lapse of time between observations, variation in conditions of observation, and variations in the observer must account for the failure of the mean r to increase according to the predicted values. So difficult to control are these factors that even at 250 minutes of observation the highest r shown is .94.

CONCLUSIONS

1. The examination of observation records of children indicates that large samples of observed behavior show a higher reliability than small samples, and the increases tend to be continually progressive, although requiring greater samples per unit of increase in reliability as perfect reliability is approached.

2. Even in fifty minute groupings, when conditions of observation are carefully controlled, and short periodic observations are made, the mean r is well within the low limit of .80 accepted by certain statisticians as a reliable coefficient, while one-hundred minute groupings are slightly below .90 and one-hundred-and-fifty minute groupings are about .92 in reliability coefficient.

3. The theoretical length of the sample for a reliability of coefficient of .95 ranges from 155 to 185 minutes when the mean r of 50 minutes is used as the base, but the actual mean r does not attain to .95 at a sample of 250 minutes, a figure greater than most studies hitherto have used.

4. The data here presented suggest that the effort necessary for ascertaining the most probable number of minutes of observation for each .01 increase in reliability in various situations would be well spent.

SEX DIFFERENCES IN HANDWRITING

PAUL THOMAS YOUNG
University of Illinois

Although most claims of graphologists are without scientific foundation,[1] there appears to be good evidence for the view that the sex of the writer can be determined from handwriting in a manner superior to chance.[2]

PREVIOUS EXPERIMENTS

Binet[3] collected 180 addressed envelopes the majority of which had passed through the mails. Of these, 89 had been addressed by women and 91 by men. This material was submitted to two expert graphologists and to 15 other persons who claimed no special skill in graphology. M. Crépieux-Jamin, one of the experts, determined the sex of the writer correctly in 78.3 per cent of the cases and M. Éloy, who examined only 103 of the series, made a percentage[4] of 75.7. The range of correct judgments for 10 of the non-expert judges was 65.9 per cent to 73 per cent. Binet concludes that it is possible, with a certain percentage of error, to determine the sex of the writer from handwriting.

Miss Downey[5] repeated Binet's experiment using 200 en-

[1] C. L. Hull and R. B. Montgomery, An Experimental Investigation of Certain Alleged Relations Between Character and Handwriting, *Psychol. Rev.*, 26, 1919, 63–74; for discussion of sex and handwriting see p. 64.

[2] J. E. Downey, Graphology and the Psychology of Handwriting, *Educ. Psychol. Monog.*, 24, 1919, 24. J. B. Watson, *Psychology from the Standpoint of a Behaviorist*, 1919, 409–11.

[3] A. Binet, Les Révélations de l'Écriture d'Après un Contrôle Scientifique, Paris, 1906, 1–22.

[4] Binet appears to have made a small error in computing percentages, or else the figures are misprinted.

[5] J. E. Downey, Judgments on the Sex of Handwriting, *Psychol. Rev.*, 17, 1910, 205–216.

velopes, 100 of which had been addressed by women and 100 by men. All but 4 had passed through the mails. As with Binet the materials were written by representative persons—teachers, business men and women, lawyers, doctors, ministers, society women and women of leisure, literary people. All were addressed to a woman to eliminate a possible cue from the sex of the person addressed. In the main series there were 13 judges all unskilled in graphology, ranging in age from 15 to something over 50 years. The range of correct judgments was from 60 per cent to 77.5 per cent, which is distinctly superior to chance and parallels the results of Binet. Miss Downey's judges indicated the degree of confidence in their judgments and for the highest degree of assurance the range was from 66.3 per cent to 83.3 per cent correct. This result confirms the finding of Binet with M. Crépieux-Jamin. Regarding the difference between masculine and feminine types of handwriting Miss Downey writes: "The typical feminine hand does appear to be colorless, conventional, neat and, usually, small." "The typical man's hand is bold or careless or experienced, above all, individual."

Newhall[6] has made a very careful study of the problem. His material was 200 mail addresses, half written by each sex without knowledge of the use to which they would be put. The writers varied in age from 20 to 60, the judges from 17 to 60 with 73 per cent in the twenties. The range of correct judgments varied from 56 per cent to 59 per cent with the mean at 57 per cent. These figures are lower than those of Binet and Downey and the range is more restricted but they are distinctly better than those to be expected by chance. Newhall prints specimens of writing most frequently judged to be masculine and most frequently judged to be feminine. Regarding the difference between the two types he writes: "In these 20 extreme cases, one is struck by the angles, irregularity and slant of the specimens judged masculine and

[6] S. M. Newhall, Sex Differences in Handwriting, *J. Appl. Psychol.*, 10, 1926, 151–161.

by the curves, regularity, and verticality of the feminine. Most of the feminine specimens, also, seem to be somewhat smaller.'' When 82 per cent of the judges agreed their judgment was objectively correct.

Kinder[7] used as material the sentence—''The dog jumps quickly over the fence after the lazy brown fox.'' This was written by 50 men and 50 women of college age. There were 20 judges, sophomores and juniors in a woman's college. The range was from 58 per cent to 76 per cent correct and the mean 68.4 per cent. Kinder prints ten samples arranged in a series from those judged correctly by 3 of the 20 judges to those judged correctly by every judge. Three samples bore the ''sex-signs'' so clearly that all judgments upon them were correct. ''These samples were neat, delicate, painstaking, and decidedly feminine.''

Broom, Thompson and Bouton[8] used as material the sentence—''Now is the time for all good men to come to the aid of the party.'' There were 40 samples selected at random, 18 written by men and 22 by women. There were 24 judges, including 8 men, 14 women, and 2 additional women who were trained penmanship teachers. Two series of judgments were obtained about 2 weeks apart. On the first series the mean for the men was 70.0 per cent and for the women 69.5 per cent correct. On the second series the mean for the men was 68.3 per cent and for the women 71.0 per cent correct. In common with other experimenters they note that ''feminine characteristics were found in the handwriting of men which caused it to be judged feminine, and masculine characteristics were found in the handwriting of women which caused it to be judged masculine.'' To illustrate 12 samples are printed. The writers state that ''regularity, curves, conventional form, and uniformity of slant are characteristic of the samples

[7] J. S. Kinder, A New Investigation of Judgments on the Sex of Handwriting, *J. Educ. Psychol.*, 17, 1926, 341–344.

[8] M. E. Broom, B. Thompson, and M. T. Bouton, Sex Differences in Handwriting, *J. Appl. Psychol.*, 13, 1929, 159–166.

judged most frequently to be feminine handwriting'' and ''irregularity, unconventional form, angles, and individual slant are characteristic of the samples judged most frequently to be masculine handwriting.''

The problem of sex differences in handwriting has been approached from directions other than graphology. Prior to the work of Binet, Awramoff[9] and Meumann[10] made instrumental studies of hand-pressure during writing, speed of movement, rhythm; and they examined the size and formation of letters. Both concluded that there are masculine and feminine types of handwriting, and that the writing of a child may be differentiated from that of an adult. An approach to the problem from the standpoint of educational psychology has been made by Starch.[11] The work of Starch indicates that girls are possibly superior to boys in the quality of handwriting and slightly superior in speed. The differences are small and their reliability an open question.

PRESENT STUDY

Experimental conditions. Specimens of handwriting were obtained from juniors and seniors in a class of general psychology at the University of Illinois. The writing was done in connection with a class assignment[12] and no one knew that the materials turned in would be used for a handwriting test. Since no restrictions were placed on the style of writing, type of pen and paper, time of writing, etc., the samples may be taken as fairly representative of the kind of writing done

[9] D. Awramoff, Arbeit und Rhythmus, *Philos. Stud.*, 18, 1903, 599, 561.

[10] E. Meumann, Vorlesungen zur Einführung in die Experimentelle Pädagogik, 1907, Bd. 2, 289–290; 1914, Bd. 3, 541f.

[11] D. Starch, The Measurement of Handwriting, *J. Educ. Psychol.*, 4, 1913, 445–464.

[12] The assignment was to match up photographs with words describing the facial expression portrayed. Several days after the assignment the written lists were turned in by 33 women and 52 men. The photographs are in F. A. Moss, *Applications of Psychology*, 1929, frontispiece.

by college students under normal, every-day conditions. From the papers handed in by the class all specimens which were written in pencil were discarded. Of the remainder a chance selection was made of 25 specimens written by men and 25 written by women. These 50 specimens were arranged in a chance order and numbered from 1 to 50. All names and clues to the sex or identity of the writer were carefully removed from the papers.

The group of judges was entirely exclusive of the group of writers. The judges were 42 juniors and seniors and 8 members of the psychology department. No judge made any pretense of skill in graphology. Of the 50 judges, 25 were men and 25 women.

In an actual experiment the samples were arranged in order from 1 to 50 and handed to the judge. He was instructed to sort the samples into two piles—in one pile those written by a man and in another pile those written by a woman. If in doubt, he was required to guess. He was told that the series contained equal numbers of M and F handwriting specimens. No limit was placed on the time or method[13] of work. The judgments were made in good daylight on a table which gave plenty of space to lay out the samples. Each judge worked alone, without on-lookers and without interruption. At the close of the assortment the judge was asked to describe the criteria he had used in judgment and to describe so far as possible the difference between M and F handwriting. A record was made of the judgment upon the sex of the writer and of the criteria used in judgment.

Results. (1) Percentage of correct judgments. Table I shows the distribution of correct judgments given by both M and F judges. The table should be read as follows—the first column indicates that 1 man and 0 woman judged correctly

[13] Some judges made an immediate decision; others made a tentative assortment, saving out doubtful cases to be decided at the end of the series.

the sex of the writer in 24 of the 50 trials, the second column indicates that 1 man and 1 woman judged correctly the sex of the writer in 25 of the 50 trials, etc.

Mere chance would give 25 correct judgments out of 50. The mean for the total group, however, is 30.5 correct judgments and the standard deviation of the mean is 2.9. The computed median is 30.25 and its standard deviation is 2.9. In terms of percentage the range of correct judgments is from 48 per cent to 72 per cent and the mean performance is 61 per cent, or 11 per cent better than chance.

TABLE I
Distribution of Correct Judgments

	NUMBER CORRECT												
	24	25	26	27	28	29	30	31	32	33	34	35	36
Male judges	1	1	1	2	2	3	3	3	0	3	2	3	1
Female judges	0	1	0	2	2	4	5	3	4	0	2	1	1
Total	1	2	1	4	4	7	8	6	4	3	4	4	2

The table shows a moderately uniform distribution of ability from zero (25 correct) or less (!) to an ability that is 22 per cent superior to chance (36 correct).[14]

(2) Sex differences in ability to judge. Our findings indicate that men and women are about equal in their ability to determine the sex of the writer from handwriting. The men made a grand total of 487 errors out of 1,250 judgments and the women a grand total of 489 errors out of 1,250 judgments.

(3) Masculine and feminine types of handwriting. The literature on the subject has a good deal to say about "sex-signs." A man may have a style of writing commonly judged to be feminine, *i.e.* having feminine "sex-signs" and a woman

[14] In the results of Binet and Downey the percentages of correct judgments were all above 60 per cent. Doubtless if larger numbers of judges had been used, the percentages would have extended down to 50 per cent.

may write in a masculine style. Consequently it is necessary to distinguish between the *apparent* sex of the writer as revealed in handwriting and the *actual* sex of the writer.[15]

To make the matter clear consider Table II. The table shows the number of times particular specimens of handwriting were *judged* to be masculine and feminine, and the *actual* sex of the writer. It will be noted that the number of judgments of masculinity plus the number of judgments of femininity always equals 50, since there were 50 judges passing on every specimen.

The second column shows a distribution of the 25 specimens actually written by men and the third column shows the distribution of the 25 specimens actually written by women. It is interesting to note that the 4 most F specimens were, in fact, all written by women and the 8 most M specimens were written by men.[16] If the results were plotted, they would appear as two overlapping distribution curves. The table shows that men tend to write in a style judged to be M and that women tend to write in a style judged to be F; but there are men whose writing appears to be F and women whose writing appears to be M.

Table II shows the difficulty involved in making a practically reliable determination of sex from handwriting. If group judgments were used, agreement of 85–100 per cent of the judges would probably indicate the sex of the writer, but in no case would this agreement be conclusive as to the actual sex.

From the theoretical standpoint interest centers on the nature of the difference between M and F types of writing. All experiments indicate that such a difference between groups of men and women exists. One of our specimens was

[15] Binet mentioned three classes of handwriting: "(1) Celles dont le sexe est très apparent; (2) Celles, moins nombreuses, dont le sexe est ambigu; (3) Celles, en petit nombre, qui portent les signes du sexe opposé."

[16] This suggests Newhall's result that when 82 per cent. of the judges agreed their judgment was correct, *op. cit.*, 161.

TABLE II

NO. TIMES JUDGED M	ACTUAL SEX OF WRITER		NO. TIMES JUDGED F
	M	F	
0		1	50
1			49
2		1	48
3			47
4		1	46
5			45
6			44
7			43
8		1	42
9	1		41
10			40
11	1	1	39
12	1	1	38
13	2		37
14	1	2	36
15	1	1	35
16	1	3	34
17		1	33
18		1	32
19		1	31
20		1	30
21			29
22	1		28
23		2	27
24		1	26
25	2	2	25
26	1	1	24
27			23
28	1		22
29			21
30	1		20
31			19
32		1	18
33			17
34			16
35			15
36	1		14
37			13
38	1		12
39	1	1	11
40		1	10
41	3		9
42			8
43			7
44			6
45	1		5
46			4
47	1		3
48	2		2
49			1
50	1		0

unanimously judged to be F and another unanimously M. These two specimens are reproduced in Plate I. Let the reader examine them and determine for himself which is M and which is F. The key is contained in the last foot-note.

(4) Criteria used. Newhall[17] collected the criteria used by his judges but did not publish the material. He writes: "The subjects reported scores of criteria of masculine and feminine writing, but as the writer knows of no way of discovering which were actually used in judging and which were rationalizations, they are not reproduced here." Despite this objection we have decided to print the adjectives and phrases used by the judges to describe M and F types:

M type	*F type*
careless (12)	careful (12)
untidy, messy (6)	neat (11)
hurried (5)	deliberate, painstaking (3)
uneven, scratchy, scrawly, irregular, slurred, halting (13)	even, uniform, precise, definite, regular (13)
letters left open (4) *t*'s not crossed, *i*'s not dotted (2)	letters completed (1)
not pretty, poor looking (3)	pretty, nice looking (3)
less legible than F (1)	more legible than M (1)
large, broad lines and letters (13) *small* (2)	small lines and letters (1) *large* (1)
heavy, firm, with force (11)	fine, delicate, dainty, light touch (12)
heavier pen (1)	finer pen point (2)
angular, daring, bold (8)	rounded, curving, flowing, symmetrical, graceful (13)
erratic, queer, individual, original (2)	conventional (1)
writing has a forward slant (5)	back-hand (13)
vertical (2)	*upright* (2)
cramped, run together (2)	*cramped*, compact, closely written, *run together* (5)
free, *spread out* (5)	*spread out* (1)

[17] *Op. cit.*, 161.

1. Grief
2. Contentment
3. Scorn (21)
4. Rage
5. Delight
6. Determination
7. Terror
8. Astonishment
9. Coquetry
10. Physical suffering
11. Bashful appeal
12. Suspicion

1. Grief
2. Contentment (44)
3. Scorn
4. Rage
5. Delight
6. Determination
7. Terror
8. Astonishment
9. Coquetry
10. Physical Suffering
11. Bashful Appeal
12. Suspicion

The above terms of description have been arranged in two columns to show so far as possible the alleged difference between M and F handwriting. The numbers in parenthesis indicate the frequency with which a particular term or characterization was used by the judges. Descriptive terms which are somewhat related in meaning have been grouped together and the number in parenthesis is the sum of their individual frequencies.

The list indicates merely the prevailing opinion of the judges. The judges are not unanimous. There are contradictions within the lists and also cases in which one and the same term is applied to both M and F types. Some of these cases have been italicized. It should be remembered that the reports of the judges lack uniformity—some are brief, trivial, casual and others are complete, considered, and well expressed. The terms listed should be taken for what they are —criteria said to have been used in distinguishing M and F types, and alleged differences between the types.

The lists do show a certain amount of uniformity. M writing is said to be careless, messy, scrawly and F writing to be careful, neat, uniform. M writing is characterized as large, broad, heavy, firm, forceful and F writing as small, fine, delicate, light, dainty. M writing is said to be angular, bold, irregular, individual, and F writing rounded, flowing, regular, conventional. M are said to write with a forward slant and F in backhand.[18]

Seven judges stated that a specimen resembled the handwriting of some M or F friend and this *similarity to a known style of writing* served as a basis for judgment.

One interesting and fairly obvious observation is that men use heavier pens than women. The difference in size of pen introduces a purely mechanical factor into the problem. There is no doubt that in an ideal experiment upon handwriting the size of pen, style of paper and kind of ink should

[18] The interested reader should refer to the descriptions and samples given by Binet and other experimenters.

be uniform; but the above lists indicate that there are basic differences between M and F types which do not depend upon size of pen.

DISCUSSION

Experiments from the time of Binet's work to the present agree that the sex of the writer may be told from handwriting in a manner superior to chance.

The interpretation of this fact offers various possibilities. The anatomist and physiologist might argue that men are stronger than women; that women have smaller hands with more fatty tissue than men; and that these physical differences express themselves directly in handwriting. The psychologist might argue that feminine characteristics such as coyness, desire to express beauty, charm and grace, etc., and masculine characteristics such as daring, strength, ruggedness, recklessness, etc., find direct expression in handwriting. The sociologist might argue that traditions have grown up regarding the general behavior fitting to boys and girls. Girls should be neat, graceful, beautiful, light of touch, uniform and rhythmical in movement, etc., whereas boys should be firm, strong, rough, rugged, angular, etc. These traditions determine M and F attitudes which to a certain extent find expression in handwriting.

It is well to realize that there may be truth in all these different types of interpretation. At present what is needed is a more thorough study of the nature of the differences in question.

CONCLUSIONS

1. Untrained judges are able to determine the sex of the writer from handwriting with an average performance 11 per cent better than chance. The range of correct judgments varies from 48 per cent to 72 per cent, whereas the result to be expected by chance is 50 per cent.

2. The present results indicate that men and women are about equal in their ability to determine the sex of the writer from specimens of handwriting.

3. There are two types of handwriting, masculine and feminine. Men tend to write the masculine type but some write a feminine or neutral hand. Women tend to write the feminine type but some write a masculine or neutral hand. The results may be visualized as two overlapping distribution curves upon a base which extends from complete masculinity to complete femininity (see Table II).

4. The criteria alleged to have been used in judgment contain some contradictions and in part they overlap. On most points there is general agreement. Masculine handwriting is characterized by such terms as careless, broad, large, forceful, angular, with forward slant; and feminine handwriting by such terms as careful, fine, delicate, rounded, with backward slant.[19]

[19] No. 21 was written by a woman and No. 44 by a man in Plate I.

A STUDY OF "WARMING-UP" IN THE CASE OF A TASK OF MORE COMPLICATED PERCEPTUAL-MOTOR COORDINATION

ERNEST B. SKAGGS
College of the City of Detroit

This paper presents the methods and results of an investigation of the so-called "warming-up" phenomenon in the case of a fairly complicated perceptual-motor task. The Goddard formboard was used, the task consisting in having the subject get all the blocks back into their proper recesses as quickly as possible. As the blocks were placed in a different (random) arrangement before each trial and as the subject could not see the arrangement until a fraction more than a second before he worked the board, the task demanded a very alert state of mind.

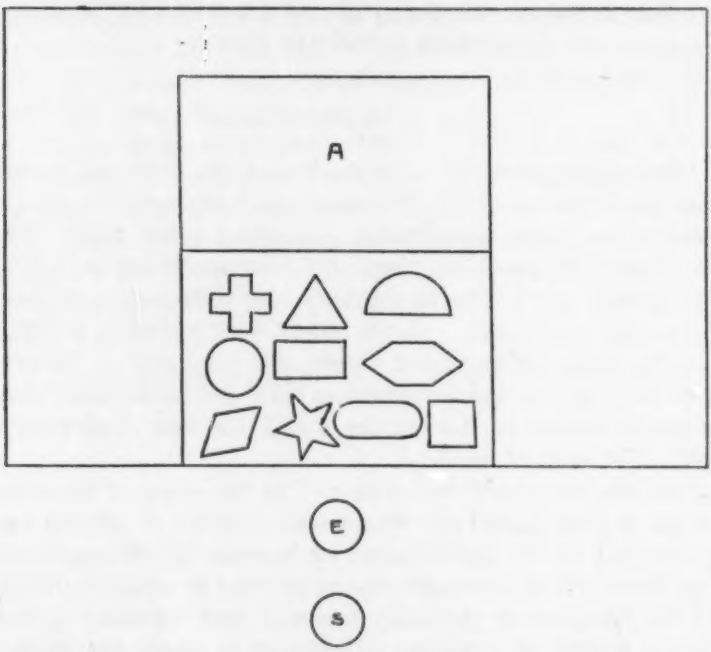
We use the term "warming-up" in the sense of the term which is now fairly well established, namely, to signify improvement in any performance or increase in efficiency over and above the improvement due to practice or actual learning. It has been rather generally assumed that warming-up has been a matter of attention adjustment in which the worker developed a more favorable attitude toward the task as he proceeded with the work. In any case where muscular adjustments or coordinations are involved it is probable that one must also take into consideration a warming-up process in the muscle tissue itself. Thus, when muscle tissue is stimulated a staircase or "Treppe" effect is found, the contractions increasing in extent for a considerable period.[1]

PROCEDURE

A Goddard formboard (standard Stoelting Company pattern) was fastened securely to the top of a table which was

[1] As discussed in the physiologies. For example, Howell's Physiology, VII Edition, page 35.

30½ inches high. A green blotter, 11½ x 19 inches, was fixed firmly to the table in front of the formboard (see A, Figure 1). At the beginning of each trial the ten blocks were mixed



in a random arrangement on the area defined by the blotter. Each block lay flat on the surface, no block was permitted to touch the edge of the board, and the blocks were separated from each other as far as the total area of the blotter permitted.

The experimenter's body formed a screen to hide the board and the arrangement of blocks from the subject before each trial. In Figure 1, E indicates the experimenter's position while S indicates the position of the subject. When all was ready the experimenter called out "Get Ready" and then stepped aside to the right, saying aloud in a slow rhythm, "One, Two, Go." As the experimenter stepped aside the

subject stepped up in position before the board. At the signal "Go," he put all the blocks into their proper recesses as quickly as possible. The experimenter started his stop-watch as he said "Go," and stopped it as he observed the last block to be properly placed. This method of procedure was very regular and constant. On the average the subject had about 1.4 seconds in which to view the board before the signal "Go" was given.

In order to have constant illumination the work was done in a dark room with artificial light. The board was illuminated by means of a 200-Watt light which was suspended 40" above the table and somewhat to the front, in order to eliminate shadows. Temperature and air movements were quite constant.

In this investigation our subjects were first given a certain amount of training in the task. Group I received only a small amount of training. Group II received a good deal more training and Group III was relatively highly trained. The meetings or trials devoted to this training will be designated as "practice-meetings" or "trials." In the practice-periods, five trials constituted a meeting. One minute of rest was taken between each trial. Ordinarily each subject met twice a week in the case of Groups I and II. The subjects of Group III met three or four times a week.

In the testing for warming-up, ten trials were taken at a meeting. One-half minute of rest was taken between each trial. In all cases, practice- or test-series, the subject was told his time after each trial.

TREATMENT OF THE DATA

Ideally, one would like to use the subjects in a task in which they had reached their absolute limit of improvement. Actually this is not possible. However highly trained a subject may be in a complicated task of the kind we have used, one can never be certain that he might not improve if he changed his method or technique. Hence we must deal with

subjects whose improvement in any experimental series may be due to practice or actual learning. We have applied the principle used by Robinson and Heron for separating true warming-up from practice effects. Unfortunately, this formula may hide actual (especially small) warming-up effects. If, upon application of the formula, warming-up is indicated, one may feel sure that it is a real phenomenon which is independent of the learning factor.

It must be remembered that in our work the smaller the time-figure the greater is the subject's efficiency. With this in mind we may indicate the following application of the Robinson-Heron formula:

Let A represent the average time for the *first five trials* at any given meeting.

Let B represent the average of the *final five trials* at this same meeting.

Let C stand for the average of the *first five trials* of the following meeting.

Now, in order to show a real warming-up, B must not only be *smaller* than A but this difference in favor of B must be *greater* than the difference between C and A, in case C is *smaller* than A. An example will make clear the situation. Let us suppose that A is 8 seconds, B is 6 seconds and C is 7 seconds. We then compute as follows: $(A-B)-(A-C)=+1$. Any positive number which we obtain thus indicates warming-up as distinct from the practice effect.

However, let us suppose that we find the following situation: A is 8 seconds, B is 6 seconds, and C is 6 seconds. Applying the formula we find that there is no warming-up indicated and the fact that B is better than A is due to learning.

In a word, then, $(A-B)-(A-C)$ must equal a positive number in order to indicate warming-up.

RESULTS OBTAINED

Group I. Subjects with a small degree of training

This investigation was carried out by the writer's assistant, Mr. Samuel Milinsky. About one month preceding the test-

ing, the writer had trained these subjects, students in his beginning psychology course, over a period of six meetings or forty-eight practice-trials. The students were now divided into two groups of equal initial formboard ability, five men and five women to a group. One group, the A Group, was definitely urged to do its best, to make a fine record. The other group, Group B, did not receive this verbal exhortation or urging before the testing.

In order to conserve space the individual tabulations of averages will be omitted in this paper. Only the statements of central tendency will be given in each case, along with statements of relative number of subjects showing or failing to show the phenomenon of "Warming-up." Any reader who is interested in these individual tables may secure them by writing to the author of this paper. A copy of these figures has been placed in the files of the library of the College of the City of Detroit.

The following summary statements may be made on the basis of the data collected with Group I:

Group A: Out of 60 records, 21, or 35 per cent, show warming-up.
The average value of warming-up is 0.30 seconds, with a mean variation of 0.209 seconds. Warming-up values range from 0.10 to 0.82 seconds.

Group B: Out of 60 records, 25, or 41.7 per cent, show warming-up.
The average value of warming-up is 0.40 seconds, with a mean variation of 0.281 seconds. Warming-up values range from 0.02 to 1.26 seconds.

From the above data we find indications, then, that a definite exhortation or urging to do one's best tends to diminish the warming-up effect. Group B, the non-urged group, shows more warming-up than Group A. On the other hand, it is evident, since the differences between the two groups are

not large, that the general experimental situation furnishes a strong incentive to "key-up" the subject to do his best.

The work with these subjects who were only slightly trained in the task indicates a positive warming-up effect in about 38 per cent of the cases, with an average time value of from 0.3 to 0.4 seconds.

Group II. Subjects with a greater amount of training

The writer carried out this investigation with a group of 8 subjects who were students in an advanced course in psychology. Three of these were women. During the training period the subjects practiced twice a week. Twenty-two practice meetings, or 110 practice trials, were obtained before the warming-up tests were given. Due to practical exigencies, only four test-meetings were possible.

Only the first three of the four meetings can be utilized or tested for warming-up as the last meeting serves as a check upon the preceding meeting. Thus, there are 24 records or cases available for study. Of these the following summary statements may be made:

(1) Out of 24 records, 11, or 45.8 per cent, show a clear warming-up effect.
(2) The average value of the warming-up time is 0.28 seconds with a mean variation of 0.18.
(3) The individual warming-up times range from 0.04 seconds to 0.90 seconds. This gives a considerable range.
(4) In 7 out of 24 cases, the second half of the work was *less efficient* than the first half. In the 6 cases where, after applying the formula, there is no warming-up shown there is still the possibility, of course, that there may be some of this effect which is obscured by the severity of the formula.

Group III. Highly trained subjects

Our third group of subjects may be designated as a highly trained group. While we may consider this group as one for

most purposes, it is convenient to sub-divide it into two groups. These sub-groups we may speak of as Group III–A and Group III–B.

Group A consisted of three very highly trained subjects, two women and the writer (subjects F, H, and G). These subjects were trained over a period of nine months, having had 93 meetings, or 465 trials.

Group B consisted of five trained subjects, A, B, C, D, and E. Subject D was a woman, the others being men. These subjects had not had quite as must training as Group A. Each had had 59 meetings or a total of 295 practice trials.

We wish to emphasize the data obtained from this group. In the first place these subjects were quite skilled at the time the warming-up tests were introduced. Their curves of progress were either very negatively accelerated or showed a plateau at the time of the testing. In the second place, these subjects were all quite well trained in psychology and can be trusted to have lived up to the conditions of the experiment as far as is humanly possible. In the third place, sufficient records were obtained to give thoroughly dependable results.

Table I presents a summary of results for all subjects after having applied the Robinson-Heron formula. Any positive value signifies warming-up effects. The term NONE signifies that the second half of the performance was actually poorer than the first half. A negative value indicates that there was no warming-up effects according to the formula.[1]

A study of the data presented in Table I permits of the following summary statements for the group as a whole:

(1) Out of a total of 113 records or meetings, 46, or 40.7 per cent, show a clear warming-up effect.
(2) The average warming-up time is 0.274 seconds with a mean variation of 0.19 seconds.

[1] The writer wishes to express his deepest obligation to these subjects who served faithfully through the year with keen interest and enthusiasm—Dr. Krueger, my colleague (C), Mr. Beckwitt (A), Mr. Milinsky (B), Mr. Gurko (E), Martha Jardon (D), Rubye Skaggs (H), and Ivah Skaggs (F).

TABLE I

Summary table for eight highly trained subjects showing results of tests for warming-up after corrections have been made for practice effects

	MEETINGS														
SUBJECT	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV
A	None	None	+.14	+.14	−.22	None	+.08	+.08	None	+.06	+.38	+.66	.0	+.34	.....
B	None	None	+.54	+.02	None	+.12	+.44	None	None	None	None	+.18	−.24	None	.....
C	+.64	+.08	None	+.42	None	None	None	+.76	−.18	+.02	+.38	+.04	None	−.34	.....
D	+.44	+.18	None	None	+.02	+.38	−.52	None	−.24	None	+.62	−.20	None	+.16	.....
E	None	−.14	None	+.68	None	None	None	+.72	None	−.78	None	None	.....	.....	.....
F	.16	+.02	−.12	+.16	+.22	+.56	+.08	+.08	−.74	None	None	None	None	+.04	None
G	None	+.40	−.08	−.14	None	None	+.34	+.20	+.52	+.16	None	None	−.28	None	−.68
H	None	None	+.10	None	None	+.20	+.38	+.24	None	None	None	+.10	−.21	None	None

(3) The warming-up time values range from 0.02 to 0.76 seconds.

When we consider the records of the individuals who compose the group the following comparisons and conclusions may be indicated:

TABLE II

SUBJECT	FREQUENCY OF W-U	AV. VALUE OF W-U
A	8 out of 14 records or 57.1%	0.235 sec.
B	5 out of 14 records or 35.7%	0.260 sec.
C	7 out of 14 records or 50.0%	0.334 sec.
D	6 out of 14 records or 42.9%	0.300 sec.
E	2 out of 12 records or 16.6%	0.700 sec.
F	8 out of 15 records or 53.3%	0.165 sec.
G	5 out of 15 records or 33.3%	0.324 sec.
H	5 out of 15 records or 33.3%	0.204 sec.

All subjects show some warming-up effects although it is evident that not all are equally susceptible to this effect. Some show much more warming-up records than others. If we compare the percentages of warming-up effects in the case of the eight subjects we note that there is a range of from 16.6 to 57.1 per cent, with an average value of 40.3 per cent. Subject E shows but little of the warming-up effect whereas subject A reveals a frequency of 57.1 per cent. Why there should be this individual difference in a group of eight highly trained, determined and interested subjects is a question which must involve a discussion of the analysis of the nature of the warming-up phenomenon as such.

If one considers only the average time-values of the records which showed a warming-up effect one notes that, with the exception of the average for subject E (0.700 sec.), the individuals are very much alike, at least the differences are small. The mean for the group is 0.246 seconds with a mean variation of 0.056, omitting the average value for subject E. This last value is omitted because it is an average of only

two records showing warming-up and is obviously an exceptional circumstance.

COMPARISON OF ALL GROUPS

Table III presents a brief summary of the results for the three groups used in this investigation, thus permitting a comparison.

TABLE III

GROUP	PER CENT OF CASES OF W-U	AVERAGE VALUE OF W-U TIMES
I—Little training:		
A, encouraged	35.0%	0.30 (M. V., 0.224)
B, non-encouraged	41.7%	0.40 (M. V., 0.293)
II—Greater degree of learning or skill	45.8%	0.28 (M. V., 0.18)
III—Highly trained	40.7%	0.27 (M. V., 0.19)

The above data offer little or no support to our original notion that the degree of learning was an important factor. Highly trained subjects, as in Group III, show just as much warming-up as subjects who are far less skilled. At least this is true if we compare the percentages of occurrence of this phenomenon. One might argue that the average time-values (see the figures at the right in Table III) indicate that there is less warming-up in the case of the highly trained group. However, Groups II and III are practically alike in time-values. This fact makes the preceding suggestion very questionable.

THE NATURE OF "WARMING-UP"

What is responsible for the initial retardation in the attainment of maximum efficiency on the part of any subject in any given task? In many cases it takes the worker some time to attain his best efficiency for any particular assignment of work. In any experimental work it is necessary to rule out

the effects of fatigue and practice. The residual effect is true warming-up. That this warming-up phenomenon appears in some cases and not in others is very significant and demands a thorough analysis of the conditioning factors.

Robinson and Heron[1] have contributed a valuable finding. These investigators say, ". . . we secured a warming-up effect separable from the practice effect and dependent for its appearance upon the relative discontinuity of the work." (Page 97.) In other words, one must control the influence of fatigue. In a later work Robinson and Bills have pointed out two factors which may operate to produce work decrement, namely, the factor of relative homogeneity-heterogeneity of the work itself and the factor of competition due to interfering factors within the work elements. Such factors as these would tend to obscure the warming-up effects.

However, we are concerned here, not so much with the conditions which obscure warming-up effects as we are concerned with an analysis of the conditions which are directly responsible for the phenomenon. Most writers have explicitly or implicitly held that warming-up is in some way related to attention adjustments. In general, this is the view which we wish to defend. Our general theses are two; first, that the warming-up is due to a process of becoming thoroughly absorbed in the task at hand and, second, that much can be done to eliminate or shorten the warming-up effects.

An examination of most out-put curves which have been obtained from workers in industry indicates a considerable warming-up. On the other hand, our data and that of most laboratory researches indicate far less of this effect. The difference is not difficult to explain in a large part at least. Our subjects stepped up to their task with a keenly competitive spirit, mentally alert and determined to do their very best right at the outset. The result is that in many cases no warming-up at all is shown and, at most, it is small in amount.

[1] See reference at end of paper.

Space does not permit of introspective reports on the part of our subjects. However, it is worth while to mention very briefly certain facts which were indicated by the eight well-trained subjects—facts which we believe go to substantiate our explanation of warming-up.

Subject A shows the greatest frequency of warming-up (57.1 per cent).

This subject always complained that he could not "get into it" in the number of trials allowed. He often came to the work with his mind full of some scientific or philosophical reading. This subject reported trouble in getting rid of his previous attitude and in becoming absorbed in the task at hand.

Subject C. This subject never reported any difficulty in getting into the task at hand but his records place him third in the list (50 per cent).

Subject E reported time and again that she always got into the task more thoroughly as it proceeded. She welcomed the shift from five to the ten trial periods.

Subjects B, G and H have about the same per cent of warming-up effects. These subjects reported no difficulty in becoming adjusted to the formboard situation. All reported that they were thoroughly keyed up and determined.

Subject E was always very alert and entered the room with a "let's go" attitude. He used a great amount of energy in working the board and never reported any trouble in getting absorbed in the work. If anything this subject was too tensed for the task.

Subject D was often fatigued and emotionally unstable. She reported much trouble in keeping her mind off extraneous and interesting events during the half-minute rest periods. Certainly this person underwent more attention adjustments during a day's session than did the other subjects.

CONCLUSIONS

1. Using a fairly complicated perceptual-motor task we have found warming-up effects in about 40 per cent of our records. We had expected to find a higher percentage.

2. The degree of skill attained seems not to be of any particular significance in our work. More skilled subjects showed as much of the warming-up effect as did relatively unskilled subjects.

3. Any condition which will make the subject keenly alert and determined to do his very best at the outset will tend to lessen the frequency of occurrence of the warming-up effect.

4. We are inclined to the view that the warming-up phenomenon is largely a matter of attitude, mental alertness and attention. It takes some time for the subject to become absorbed in his task and to disengage his mind from other matters.

5. Stated in purely physiological terms we may say that the warming-up effect represents, for the most part, a phenomenon of the central nervous system. The locus of the phenomenon in the neuro-muscular system is, we believe, in the cortical areas, involving a proper arousal of cortical "sets" or the exercise of nerve-synapse systems.

REFERENCES

1. Arai, T. Mental Fatigue. Columbia University Contributions to Education, number 54, 1912.
2. Batty, A. M. Some Observations upon Practice and Fatigue as They Affect the Rate of Tapping. American Journal of Psychology, 1909, 20, 449–455.
3. Heron, W. T. The Warming-Up Effect in Learning Nonsense Syllables. Journal of Genetic Psychology, 1928, 35, 219–228.
4. Robinson, E. S., and Heron, W. T. The Warming-Up Effect. Journal of Experimental Psychology, 1924, 7, 81–97.
5. Robinson, E. S., and Bills, A. G. Two Factors in the Work Decrement. Journal of Experimental Psychology, 1926, 9, 415–443.
6. Thorndike, E. L. Educational Psychology. Briefer Course. Chapter XIX, pages 283–329.
7. Watson, J. B. Psychology. 1924. Pages 354–355.
8. Wells, F. L. Normal Performance in the Tapping Test. American Journal of Psychology, 1908, 19, 437–483.
9. Wimms, J. H. The Relative Effects of Fatigue and Practice Produced by Different Kinds of Mental Work. British Journal of Psychology, 1906–1908, 2, 153–195.

STUDENTS' REACTIONS TO A FIRST COURSE IN PSYCHOLOGY

GEORGENE HOFFMAN SEWARD

Barnard College

PROBLEM

Although the blanket description in the college catalog gives an impression of uniformity to the uninitiated, a first course in psychology may have as many connotations as there are teachers and textbooks. According to some, the introductory course should lay the foundation for more advanced work, acquainting the student with the theoretical questions, and at the same time providing laboratory experience in the scientific handling and statistical treatment of problems interesting to the professional psychologist. Thus Skaggs (9) comments on the first course in psychology: "I should like to see the day come when the American Psychological Association will go on record as recommending that no scientific credit be given in psychology unless the laboratory work accompanies the textbook course." A contrasting viewpoint adopted by other members of the profession regards the elementary course as a general survey of the field, with the object of giving the student an airplane view of the science, showing its outlines and surrounding regions.

A suggested solution to this disagreement among the doctors is to offer two varieties of beginners' psychology—one for students who intend to continue their work in the subject; the other for those who do not plan further study along this line. The obvious criticism of this compromise is that the students are seldom sufficiently familiar with the nature of psychology to decide beforehand whether or not they shall go beyond the first course, and, in fact, usually make such decisions as a result of this course. Therefore the students' re-

actions to their initial contact with the subject cannot be ignored and offer a suggestive approach to the construction of a first course in psychology. The wise student, profiting by the experience of his upperclassmen, registers with the teacher whose reputed treatment of the subject gives the most promise of satisfying his interests and needs. While it is necessary to disillusion him as regards certain popular confusions of psychology with various forms of fortune telling and occultism, the course is bound to defeat its own end if, by a too academic emphasis it creates an attitude of disappointment, boredom or derision. As Schoen (8) says, "Psychology should be the most vital course in the curriculum." The present study is an experimental analysis of a first course in the hope of better meeting the students' needs.

PREVIOUS STUDIES

Several studies reported in the literature bear upon the present problem, and in spite of differences in technic, offer a basis of comparison with the present results. In 1923 Laird (5) took students' votes on topics of greatest interest and value in a "usual course" in elementary psychology, "well over one third of which was devoted to sensations and sense organs, and the remainder to structural psychology of the thinking-feeling-doing, sensation-perception-conception-judgment-reasoning type." In the returns, the category, Abnormal Psychology, led with an interest vote of 56 out of 120, and Applied Psychology came in second. The topics at the opposite extreme which received no votes for interest were Mental Structure, and Animal Behavior. On the value ballot, Applied and Abnormal Psychology exchanged places. The correlation between interest and value was found to be + .33. In another vote at the end of a revised course in which the last six weeks were devoted to mental hygiene, and in which applied and abnormal psychology were emphasized, Laird's returns show 68 and 67 votes out of 70 for Mental Hygiene in interest and value respectively. This result is hardly sur-

prising nor is it especially significant in view of the overweighting which the topic received in class treatment. Only with the caution of these considerations may we accept the author's conclusions that "college students are eager for instruction in mental hygiene especially as it affects them personally. It was estimated to be of more value than any other part of the course in elementary psychology."

In 1926 Arnold (1) required 300 students to vote on the interest and value of various divisions of the lecture and textbook topics in a first course in psychology. The subjects were not limited to casting votes for the topics of *greatest* interest and value as in the case of Laird's investigation, but were free to check every topic in which they felt special interest, and every one in which they recognized value regardless of interest. Since the lecture topics have few counterparts in the present study, comparisons will be confined to the textbook divisions. Of these, the ones rating highest for interest were Dreams, Habit, Instinct, Emotions, and Play, while the lowest positions were held by Esthetics, Neural Theory, Mind-and-Body, Perception and Nervous System. The value votes were somewhat different, showing a rise in popularity for the physiological topics. The correlation between the two orders was reported as + .60.

During the same year G. B. Watson (10) made a questionary study of educational psychology to determine the practical value of the course for education rather than for the students' personal benefit. He used 400 subjects whose experience or interest was in the field of education. To them he presented in question form 15 groups of problems covered in a course in educational psychology, requesting them to assign numerical ratings. He found that "the apparent concurrence of opinion would place major emphasis upon problems of emotion and personality adjustment, problems of original nature and heredity, and problems of general teaching methods. Little emphasis would be placed upon the psychology of the acquisition of skill, or the psychological prob-

lems involved in home-making and pre-school child training, or the various schools and theories in psychology."

Hertzberg (4) in 1928, making an extension of Watson's study with the same problems, had the faculty of a teacher-training institution rate them with regard to their value to the elementary school teacher. The results were confirmatory of those of the earlier research, personality adjustment problems ranking first; schools and theories of psychology, last.

Another experiment in 1928 was performed by Lauer (6). An orthodox basic course in general psychology was analyzed to discover what the students felt would be useful to them in their future work; 108 students served as subjects. They were presented with a questionary covering five main divisions of the course, and their corresponding subdivisions. The students were allowed three choices in determining each of the following—which of the headings they considered most helpful, which might have been expanded, and which might have been abbreviated. Relative frequencies of suggested topics having vocational value showed that "these students favor a study of mental life as such, and do not appreciate the physiological connections necessary to a basic knowledge of the working of the mind." They desired elaborations concerning heredity and environment, memory and other conscious processes, learning, control of emotions, and the influence of personality in school and business. On the other hand, they felt that overemphasis had been laid upon ontogenetic and phylogenetic development of the organism, nervous system and physiological psychology, reflex action and instincts.

METHOD

In the present study 87 Hunter College juniors, students in the writer's one-semester course in elementary psychology, served as subjects in the experiment. A series of topics treated during the term was listed in random order. In an effort to give approximately equal weight to all the topics, they were so chosen that each represented three lecture

periods. Language, however, which received only one class hour was included along with the others because the writer desired to know whether it proved sufficiently valuable to the students to warrant further expansion. The list of topics was presented to the students on the blackboard during their last week of recitations. In order to avoid the influence of inequalities in the recall of the various topics, a brief summary of the chief subdivisions included within each category was given orally by the experimenter. The series with the subheadings in the order presented to the subject is as follows:

1. *Imagery and Association*—kinds of imagery; imageless thought controversy; theories of thinking; free and controlled association, from dreaming to reasoning.
2. *Decision, Mental Conflict, and Mental Adjustment*—voluntary activity; reaction time; mental mechanisms; integration.
4. *Physiological Substrata*—central nervous system; autonomic nervous system; endocrines; cerebral localization.
5. *Instinct*—heredity and environment; instinct controversy; dependable motives.
6. *Character and Personality*—measurement by ratings, questionaries, experiments, genetic and clinical studies; methods of psychotherapy; personality "types."
7. *Perception*—space, including theories of sensory organization, and third dimension; time; motion; illusions.
8. *Motor Learning*—survey of animal learning; theories of learning; rôle of motivation.
9. *Language*—development in the race—gesture, oral and written; development in the individual.
10. *Intelligence*—survey of tests; rôle of heredity and environment; theories of intelligence; special abilities and disabilities.
11. *Emotion*—physiological expression, facial and vocal expression; theories of emotion; feelings.
12. *History and Contemporary Schools*—contributions from physiology, biology and medicine to the development of psychology; survey of the contemporary schools of introspectionism, behaviorism, Gestalt, purposivism and psychoanalysis.
13. *Sensation*—mechanism of vision, hearing, taste, smell, skin and proprioceptive sensations.

The subjects were instructed to copy the list of main headings in the arrangement given, and then to rank the topics

first in the order of their preference, and secondly in the order of the value they had derived from them. Since value received depends upon the scale by which it is measured, it was important to determine the various criteria by which the students evaluated the course. Accordingly they were asked to state the predominant standard by which they judged the topics. Owing to limitation of time we were unable to have the rankings for preference and value made on separate days. The subjects, however, were given special instruction in making the value rankings, to disregard the previous preference rankings as far as possible. To insure frankness in these estimations, all papers were returned to the experimenter without signatures.

RESULTS AND INTERPRETATIONS

Table IA shows the order of relative preference of the topics according to the average rankings of the subjects. Table IB shows the order of value without regard to the specific criteria used. An analysis of the influence these criteria exert on the value judgments will be presented later.

It may be observed from Table IA that the highest positions for preference are given to Character, Intelligence, Decision, and Emotions; the lowest to History, Perception, Sensation, and Physiological Substrata. On the value scale there are no outstanding shifts in rank, the upper and lower halves of the series retaining the same topics.

Although the "raw" rankings shown in Table I are interesting, their quantitative significance can be determined only from the statistical reliabilities of the differences between pairs of topics. In Table II are presented the ratios, $(\frac{D}{\sigma_D})$, for all permutations and combinations of the divisions of the course. When a ratio of three or more is found the difference is considered "significant," and in cases where it is practically certain to exceed this figure the ratio has not been computed but is represented in the table by dots.

TABLE I

A–PREFERENCES

Topic	Average	σ(dist.)	σ(av.)
Char.	3.53	3.05	0.33
Intell.	3.84	2.57	0.28
Dec.	5.28	3.55	0.38
Emot.	5.45	2.62	0.28
Imagy.	5.62	3.35	0.36
Mem.	6.31	2.27	0.24
Inst.	6.62	2.87	0.31
Lang.	7.44	2.94	0.31
Mot. L.	8.12	3.20	0.34
Hist.	8.86	3.43	0.37
Perc.	9.44	2.82	0.30
Sens.	9.75	2.61	0.28
Phys.	11.00	3.03	0.33

B–VALUES

Topic	Average	σ(dist.)	σ(av.)
Intell.	4.34	2.83	0.30
Dec.	4.70	3.61	0.39
Char.	4.78	3.45	0.37
Inst.	5.75	3.06	0.33
Emot.	5.88	3.24	0.35
Mem.	6.00	2.91	0.31
Imagy.	6.21	3.05	0.33
Mot. L.	7.58	3.18	0.34
Phys.	8.42	4.16	0.45
Perc.	8.55	2.86	0.31
Sens.	8.86	3.12	0.33
Hist.	9.17	3.86	0.41
Lang.	9.35	2.91	0.31

An examination of Table IIA indicates that not only is the topic Character given the greatest preference but that its only rival is Intelligence, since together they share honors as significantly higher than all other topics in the course. Decision, Emotion and Imagery do not show dependable differences from one another but are all significantly higher in preference ranking than Language and those lower in the scale. Memory and Instinct, though not reliably exceeding Language, are significantly above Motor Learning and those below. Language in turn represents a "real" difference from Perception, which is about on a par with Sensation. There

TABLE II. RELIABILITIES OF DIFFERENCES $(\frac{D}{D_\sigma})$

A. PREFERENCES

	Char.	Intell.	Dec.	Emot.	Imagy.	Mem.	Inst.	Lang.	Mot.L.	Hist.	Perc.	Sens.
Intell.	0.70											
Dec.	3.50	3.06										
Emot.			0.36									
Imagy.			0.65	0.37								
Mem.			2.29	2.32	1.57							
Inst.			2.73	2.79	2.08	0.78						
Lang.			4.41	4.74	3.79	2.82	1.82					
Mot.L.						4.31	3.19	1.45				
Hist.								2.90	1.45			
Perc.								4.66	2.87	1.21		
Sens.									3.62	1.90	0.51	
Phys.										4.28	3.47	2.82

B. VALUES

	Intell.	Dec.	Char.	Inst.	Emot.	Mem.	Imagy.	Mot.L.	Phys.	Perc.	Sens.	Hist.
Dec.	0.73											
Char.	0.92	0.15										
Inst.	3.13	2.06	1.94									
Emot.		2.27	2.16	0.27								
Mem.		2.60	2.49	0.54	0.26							
Imagy.		2.96	2.86	0.98	0.69	0.46						
Mot.L.		5.54	5.49	3.81	3.47	3.15	2.85					
Phys.							3.95	1.47				
Perc.								2.06	0.24			
Sens.								2.67	0.79	0.67		
Hist.								2.94	1.23	1.19	0.66	
Lang.								3.77	1.69	1.78	1.07	0.35

are no significant differences among History, Perception and Sensation. Physiological Substrata, however, which is least desirable according to students' ratings, is statistically entitled to this position.

Upon consideration of Table IIB, Intelligence is found to lead all except Decision and Character, which do not significantly differ from each other. The array of topics is sharply divided at the point of Motor Learning, which is given a reliably lower evaluation than any of those which precede, with the possible exception of Imagery. The only other significant differences are between Motor Learning and Language

and between Physiological Substrata and Imagery. Ratios which most closely approximate statistical significance are those between Motor Learning on the one hand and Perception, Sensation and History on the other. In general, Motor Learning seems to divide the series into two fairly compact groups.

Summarizing the preceding results, we find that the topics Intelligence, Decision and Character, all of which have practical bearing on the personal lives of the students, were rated at the uppermost end of both the preference and value series. Similar trends were noted in the findings of previous investigators. Confirmation of earlier studies is also apparent at the opposite end of the series. On both scales History is among the less desirable topics. This result, which was also reported by Watson and Hertzberg, may be largely due to the inadequacy of the students' background. Language is another subject which makes a poor showing in the course and might well be omitted to give room for a more favorable topic. Physiological Substrata also appears in the lower part of the scale. Its rise from thirteenth place in preference rankings to ninth in value is more apparent than real as there are no significant differences between this topic and any of those below it. Sensation and Perception, which have a large physiological component, receive low ratings in both series. These results point in the same direction as those of Arnold and Lauer and probably reflect the students' resistance to physiological topics. This attitude can be easily understood when we consider that in a one-semester course so much material with direct psychological import must be sacrificed to establish a background of physiological information whose value for correlation with psychological function must be questionable in these days of changing neurological theory. Until some of these hypotheses have been more firmly established in the laboratory of the neurologist or physiological psychologist, knowledge along this line cannot be used as Meyer (7) suggests "to explain with the most rigorous logic

of the mathematician the rôle played by the nervous system in this transformation of the infant into the adult."

The analysis of the results would not be complete without determining the relationship between preference and value rankings. This is expressed by a correlation of $+.81 \pm .06$, obtained by the Spearman rank-difference method. The discrepancy between this result and that of Laird who reports a corresponding coefficient of $+.33$ may possibly be explained in terms of differences in technic. Laird limited his subjects to voting for the topics which possessed the highest degrees of interest and value. According to this method there are probably fewer chances that the two series would correspond than by the order of merit method used in the present study. Arnold, who used the voting technic but gave his subjects more latitude than Laird in their selections, found a correlation of $+.60$ between interest and value. This is considerably nearer the present figure.

A comparison of preference and value rankings in Table I indicates a wider range of averages in the case of preferences. This suggests a higher degree of consistency in the students' preference reactions than in their value reactions, an impression borne out by a comparison of the averages of the σ(dist)'s for preferences and values, which are respectively 2.95 and 3.25. The difference between these averages is 2.73 times the σ(diff). Although this does not represent complete statistical reliability, the chances that the obtained difference is a "true" one are found to be 9968 in 10,000.

The lesser consistency in the value rankings may possibly be explained in terms of the varying criteria used by different subjects in making their evaluation. A further analysis of the value ratings was made according to the separate criteria used, which are as follows:

The degree of correspondence among these criteria may be determined from the intercorrelations presented in Table III. In these comparisons only the first four criteria were included. The others were omitted because of the small number of subjects using them. In this connection it is interesting to note

Criterion	*Number of Students*
Curiosity	30
Knowledge of Human Nature	15
Vocational Suggestions	14
Personal Problems	13
General Information	7
Academic Credits	6
No Criterion Stated	2
Total	87

that although the course is required of all, only 7 per cent used Academic Credits as their criterion of value. This is in surprising contrast to results reported by Book (2) according to which 66 per cent of the students in a course in educational psychology took it because it was required.

TABLE III
Intercorrelations among Criteria of Value*

	VOC.	HUM. NAT.	PERS. PROB.
Cur.	+.80	+.78	+.76
Voc.		+.95	+.72
Hum. Nat.			+.77

* Spearman rank difference method.

Examination of the high positive correlations in Table III indicates considerable overlapping among the various criteria. Although various interpretations may be given to these findings, they may simply show that similar factors determined the students' evaluations of the topics when the criteria were differently expressed.

A more detailed comparison of the rankings assigned the different topics under different standards of evaluation is suggestive of certain interesting variations, but statistical significance has not been determined. For example, Instinct is ranked second according to the standard of Curiosity, while it rates only eighth when the vocational criterion is used. This is also true when Curiosity is compared with Knowledge of Human Nature as a means of evaluating the topics. In

the latter comparison, moreover, Decision, which ranks first for Knowledge of Human Nature, receives a position of 5.5 for Curiosity. This same relationship holds when Curiosity ranks are compared with those of Personal Problems. Still another comparison shows that Character, which received a factual and academic treatment, ranks second for Vocational Suggestions but drops to seventh for Personal Problems.

Such comparisons as these seem to indicate certain differences in the value which various topics attain according to the special needs of the students and point out the possibility of better adapting the first course to meet these varying needs, provided they are determined beforehand. English (3) in a study of "why students register for psychology," found the vocational motive the most clearly crystallized, occurring in 74 per cent of students; 42 per cent "rather hope that they will learn to understand themselves and others better, but this is hardly specific." Twenty-five per cent want to learn about the mind. Further investigations along this line would indicate subjects on which special emphasis should be placed in the elementary course in psychology and would be of great assistance to the teacher in framing such a course.

SUMMARY AND CONCLUSIONS

The present study represents an attempt to analyze a first course in psychology according to the topics discussed. We presented a list of the topics selected so as to be approximately equal in the amount of class time spent on them, to a group of 87 women students. The subjects were instructed to rank the topics in the orders of preference and value. The statistical reliabilities of the differences in ratings were determined. It was found that:

1. The highest ranks for both preference and value were assigned to Intelligence, Character, and Decision. These are obviously the topics of the greatest practical importance for personal adjustment.

2. The topics of least preference were Perception, Sensation and Physiological Substrata, and those of least value were Sensation, History and Language.

3. The correlation between preference rankings and value rankings was $+.81 \pm .06$.

4. Intercorrelations of value when different criteria were used were all over $+.70$.

The results of this study offer an approach toward adapting the first course in psychology to better meet the needs of the students whom it is expected to benefit. Within the limits of this investigation at least, there is justification for greater emphasis upon those topics which will lead toward personal adjustment, and less emphasis upon perception, sensation, physiological substrata, and history and contemporary schools. Moreover, it would be desirable, in organizing the course, to make a preliminary survey of the students' motives for taking it so that these may be satisfied as far as it is deemed feasible.

BIBLIOGRAPHY

1. Arnold, H. J. What parts of elementary psychology are most interesting to students? Ped. Sem. 1926, 33, 729–35.
2. Book, W. F. How to develop an interest in one's tasks and work. J. Educ. Psychol. 1927, 18, 1–10.
3. English, H. B. Why students register for psychology. J. Appl. Psychol. 1928, 12, 242–4.
4. Hertzberg, O. E. The opinion of a teacher-training institution concerning the relative value of subject-matter in educational psychology to the elementary school teacher. J. Educ. Psychol. 1928, 19, 329–42.
5. Laird, D. A. The reaction of college students to mental hygiene. M. Hy. 1923, 7, 271–6.
6. Lauer, A. R. Some aspects to be considered in the organization of a first course in psychology. Psychol. Monog. 1928, 39, 110–18.
7. Meyer, M. Most recent text-books of psychology. Psychol. Bull. 1927, 24, 359–77.
8. Schoen, M. The elementary courses in psychology. Amer. J. Psychol. 1926, 37, 593–9.
9. Skaggs, E. B. Dr. Schoen on "The elementary courses in psychology." Amer. J. Psychol. (notes) 1927, 38, 153–4.
10. Watson, G. B. What shall be taught in educational psychology? J. Educ. Psychol. 1926, 17, 577–96.

NOTES AND NEWS

Silver Anniversary Celebrated by Vineland Laboratory

The Training School at Vineland celebrated the 25th anniversary of its Research Department September 3–5, by a series of meetings which were attended by more than fifty leaders in the fields of education, psychology, sociology and medicine. The first Director of the Department, Dr. Henry H. Goddard, was present as guest of honor.

The important contributions of this laboratory since its inception were reviewed. With regard to research now under way, Dr. Edgar A. Doll, present Director of the Department, told of studies of unusual mental growth, of success on parole for feeble-minded, an etiological survey, extensive research on Idiocy and Mongolism, and the development of the gait graph and motion picture technique in furthering the study of birth injury as a category of mental deficiency. Dr. Doll exhibited several motion pictures of children under muscle training treatment. "Any one of these children," he suggested, "is comparable to Helen Keller. These children are just as fascinating from all points of view; just as many possibilities are there."

With regard to the important findings of this and other child study centers, E. R. Johnstone, Director of The Training School, expressed doubt as to whether the discoveries of the clinics are reaching teachers, or whether they are "too frequently considered too sacred for human hands or too technical for ordinary intelligent minds, and so find repose in the files of the clinic or of the principal?" He expressed the opinion that testers and examiners must provide the teacher with much more information than she ordinarily gets, if she is really to prepare the child for the future.

"Today," he said, "the physician, the psychiatrist, the psychologist and the school man all unite to tell us where we may aim to take the delinquents or those who are mentally deficient. Shall we not give the normal child the same advantages?

"How far ahead shall the teacher look? Does that mean how far today, or this year? Or shall she be given all of the scientific information possible, so that she may envision him five or ten or even twenty years hence? May we not look forward to a much better understanding of a child's aptitudes and his possibilities? Every child who misses making a grade is a challenge to us. Every square peg in a round hole is an example of our ignorance."

Dr. Arnold Gesell, Director of the Yale Psycho-Clinic, presented moving pictures, showing the formboard reactions of a normal infant, from the age of 20 weeks to 2 years. As pointed out by Dr. Gesell, this record "reveals a steady, consistent development of progression which has even mathematical suggestions which indicate that development, even in the complicated field of mental growth, is not whimsy but is fundamentally lawful. It is the developmental increase in the elaboration in behavior patterns, which is, indeed, so lawful that even the apparently trivial behavior of the infant in relation to the formboard will repay systematic study."

The addresses, reports and discussions will appear in the near future in a special commemorative volume. Those interested in securing a copy of the printed proceedings should address the Vineland Training School, Vineland, N. J.

Teachers College, Columbia University, announces a plan for a new type of teacher training institution to open September, 1932. This new college, an undergraduate unit at the college level, will attempt to demonstrate radically different methods in the selection and training of young men and women who are to become teachers in nursery, elementary and secondary schools. Rigid methods of elimination will be used in the selection of students. Selection will be based on good health, sound scholarship, desirable personal qualities and promise of unusual growth. The student body, limited the first year to 100 young men and 100 young women, will not be restricted to local sections of the country. The college will reject the traditional point system. Graduation will be based upon satisfactory examinations of practical as well as academic character. The unusual care in creating the student body will be taken because it is felt that mediocrity is today the curse of the teaching profession.

The National Child Labor Committee announces that Child Labor Sunday will be observed for the twenty-fifth year on January 24, 1932. It stresses the fact that cutting short children's schooling for work benefits no one, tends to depress wage rates, lowers standards of living, and aggravates unemployment. The Committee offers programs, literature, and posters for use on Child Labor Sunday to those who wish to bring these facts before the people through churches, schools, clubs and forums. Those interested may address the National Child Labor Committee, 331 Fourth Avenue, New York City.

A survey of fifty-two land grant colleges and universities made by the Federal Office of Education recently indicated that a large propor-

tion of students in the United States earn their way through college. The study revealed that 24,952 students out of 34,589 alumni and former students, or 72 per cent, were self-supporting wholly or in part while they attended college. These land grant colleges, established to provide higher educational opportunities for those who could not afford an education, are supported by public taxation. A greater proportion of students in agricultural courses were self-supporting than those in any other field of study, the percentage being 86. Students in engineering and those preparing to teach were next on the list. Only half of the women studying home economics were wholly or partially self-supporting due to the fact that opportunities for women to earn their way were much fewer than for the men.

Social Science Research Council offers for 1932–33 Grants-in-Aid of Research in the Social Sciences to mature scholars, without reference to age, whose capacity to carry on productive research has been clearly demonstrated. These include grants in economics, social, economic and political history, political science, social psychology, sociology, cultural anthropology, statistics, law and human geography. The grants were established by the Council with a view to assisting members of staffs of institutions which cannot provide funds for social science research. The maximum amount granted will ordinarily not exceed $1,000. The closing date for applications for 1932–33 is February 1, 1932. For further information those interested may address the Grants-in-Aid Secretary, Social Science Research Council, 230 Park Avenue, New York City.

The Social Science Research Council also offers a number of research fellowships in the social sciences. These fellowships are open to men and women who possess the Ph.D. degree, or its equivalent in training and experience, and who are not over thirty-five years of age. The basic stipend for a period of twelve months is $1,800 for single Fellows and $2,500 for married Fellows. Closing date for application for 1932–33 is December 1, 1931. Further information may be obtained at the address given above.

Dr. Herman M. Adler, until recently the Director and Secretary of the Behavior Research Fund of Chicago, goes to his new field with the assurance of the esteem and affection of his co-workers in the Fund. At the last meeting over which Dr. Adler presided in an official way the Trustees and Staff of the Research Fund passed resolutions in which they expressed their appreciation of his long services among them. "For the staff whose privilege it has been to serve with Dr. Herman M. Adler," says Dr. Karl S. Lashley, "his leadership has been one of courage and inspiration. With vision to see the problems of human welfare

in their broadest relations and with clear judgment of values and practical means he has symbolized for us both a high ideal and the hope of attaining it. . . . Our message to Dr. Adler is one of deep personal affection, of appreciation of his stimulating guidance and scientific insight, and of continued faith in the ideals for which he stands.''

Dr. Adolph Meyer says in part, ''Behavior Research and Chicago and Herman Adler are so intimately identified with Psychiatry and Mental Hygiene that Adler's going to another field is of far more than local concern. As a psychiatrist and hygienist, Adler entered a sphere prepared by William Healy and he leaves it richly expanded and with a stamp and record of achievement and vitality expressing as nothing else could do the vision and effectiveness of a real leader and generator of practical and research work in criminology and behavior in this country.''

A recent study by Dr. James F. Rogers, Federal Office of Education Health Specialist, revealed that Binet tests of children with speech defects in schools of St. Paul, Minnesota, show that such children were neither duller nor brighter than other children, and it is only because the stammering or stuttering child cannot display normally his native ability that he has often been considered of low mentality. The study also showed that speech defects are at least twice as common among boys as girls and they occur about twice as frequently in the negro as in the white race. A survey revealed that 80 school systems now employ special teachers to give attention to speech-defective boys and girls.

A new test of knowledge of social standards entitled ''The Best Thing To Do'' has recently been published by the author, Frank E. Tomlin, and the Stanford University Press. This test is designed for use in connection with character-education programs in the elementary grades of the public school. It is of the multiple-answer type and consists of forty-five situations in which the pupil is asked to designate the best and worst of four possible answers. Norms are now established for the fourth, fifth and sixth grades, and tentatively for the seventh and eighth.

BOOK REVIEWS

GARDNER MURPHY AND LOIS BARCLAY MURPHY. *Experimental Social Psychology.* Harper and Brothers, New York. 1931. 709 pp. $3.50.

Writers of books summarizing literature often have a thankless task in spite of the fact that their work is of greatest importance. Such writing seems to fall along a scale extending from the least valuable, *i.e.*, a mere recital of experimental work following a "Resolution—whereas—item 1—item 2—etc." sort of technic, to the extremely significant work of critical interpretation. It is a book of the latter type that the Murphys have succeeded in writing. Without doubt this work will establish itself as a guide to the region of science devoted to social psychology.

The fundamental view-point is well expressed in a sentence from page 488: "It must, of course, be borne in mind that here—as in so many other places in this book—our interest is to open experimental questions, not to close them." For this reviewer, at least, this interest has been kept foremost; in reading he has been struck with the large number of questions raised which are amenable to experimental treatment. Any graduate student who is at all interested in the social aspects of psychology, and has not yet found a problem, should carefully read this book. If he cannot find a dissertation topic, or two, he ought to retire to the simpler things of life.

In attempting to present an integrated picture of the state of the science at present the authors have followed an organization not usually found in text-books on social psychology. Part one, devoted to "Basic Principles," introduces the reader to some basic biological and psychological concepts dealing with motives, nature and nurture, individual differences and learning, as well as orienting social psychology among the social sciences. Part two, "A Genetic Study of Social Behavior," describes methods of studying social behavior in children and their application from infancy to adolescence. Part three, concerned with the "General Laws of Social Interaction in Our Own Society," discusses the behavior of individuals in the group situation, the behavior of cooperating groups, and treats the measurement of personality and attitudes in the adult. Each chapter follows the preceding one logically, while the later chapters are related to the earlier ones, thus making an admirably integrated whole.

The authors say of their book (p. 9), "Its business is grubby, mole-like; it digs here and there in a thousand interesting places accessible by means of its own instruments; it aims at the faithful presentation and interpretation of specific methods and discoveries characteristic of the experimental and quantitative social psychology of recent years." The "here and there" into which this grubbing has been done is documented by 822 notes appended to the chapters. And these give evidence that they are critical selections from a much greater amount of material. Such intelligent industry, like virtue, is its own reward; yet, like virtue, it has many further compensations.

In closing it must be said that the House of Harpers has mightily strengthened its psychology list with this title and has redeemed itself for past errors.

L. L. THURSTONE AND RICHARD L. JENKINS. *Order of Birth, Parent-age, and Intelligence.* University of Chicago Press, Chicago. 1931. xiii + 135 pp. $3.50.

This monograph of the Behavior Research Fund consists of two parts. The first is a statistical study of about 1,500 cases drawn from the records of the Institute for Juvenile Research, for the purpose of discovering what relations exist between birth order and intelligence and between parents' age and intelligence. The second part is an extensive survey of the literature on the relation of birth order to various anthropological, psychological and medical data.

The raw data from the 1,500 cases indicate a decrease in mean I.Q. with increased ordinal number of birth. However, family size and socio-economic status may both operate to bring about this negative correlation. When these factors are controlled by comparing only siblings and weighting the differences between the various ordinal number pairings (which method is illustrated statistically) the correlation is reversed. In only five of twenty-eight possible pairings does the earlier born have the higher mean I.Q. "The later born children seem to be brighter on the average than their earlier born siblings, and the variability of test intelligence seems to increase with order of birth."

The same data serve for the study of the relation of intelligence to such factors as fecundity of parents, miscarriages, sex of siblings, twinning, age of parents, time intervals between births, and age of patients. These are all summarized in eleven points which, in spite of their interest, we cannot reproduce in full. Among the more important findings we might mention that with dull, adequate and superior children a negative correlation is evident between I.Q. and size of sibship. Alternation of sex in birth order indicates a better prognosis for behavioral adjustment. There is no indication that the age of either parent affects intel-

ligence, although parents of nearly the same age tend to have more intelligent children.

The second part, chiefly the work of Dr. Jenkins, is a non-critical summary of the literature related to the problems of birth order and parents' age, not only in relation to intelligence but also to various diseases, anthropometric measures, crime, etc. The bibliography includes 175 references giving evidence that the literature has been well reviewed. The text of the review with its multiplicity of tables will prove extremely valuable.

This monograph is an important contribution to a practically and scientifically important problem. The reviewer's only criticism is that one or two sections of Part I and several sections of Part II have the appearance of "space-fillers." Still this addition of material may be a virtue, not a vice.

C. M. Louttit,
Indiana University.

Boven, W. *La Science du Charactère.* Paris. Delachaux et Niestlé, 1931, pp. 351.

Character is defined as the psychology of an individual. This broad definition makes almost any psychological topic pertinent to the discussion. The early chapters deal with the nature or structure of character. Dispositions are correlated with the vegetative motor and sensory systems; traits with a hypothetical coenaesthetic brain center; and features (linéaments) with the reasoning process. These chapters are summarized by a sort of form or questionnaire for studying the character of any individual. In this questionnaire the first consideration is physiological; the second involves traits,—(1) attitudes toward the bodily self such as eating and sex, (2) attitudes toward one's own personality such as pessimism and fear, (3) attitudes toward others such as dominance and obstinacy, (4) attitudes toward things such as love of nature; the third consideration includes types of reflection, fields of interest and attitude toward the abstract. This outline might be of some interest to the clinical psychologist as suggesting items to include in personal history blanks.

Intelligence is described as an ensemble of very closely coordinated centers provided with an energy in the form of attention. Its integrative aspect is stressed and Binet tests are described. Types of intelligence such as objective and subjective are mentioned but nothing is said about the common classification of abstract, mechanical and social.

The chapter on measurement of character is the one that is presumably of most interest to the experimentalist. The difficulty is pointed out that one can study intelligence "cold" but character must be

studied "hot." The author feels that it is not very profitable to split it up into elements inasmuch as courage, for example, may involve many different emotions. However he classifies methods as: observation of character in life; self analysis; morphology and physiology; graphology; tests. Quite a few test efforts are described although the author feels that their value is rather slight and that they give us little more than we get by intuitive methods. He makes one astounding statement to the effect that graphology is "one of the best measurements of character, one of the most practical and sure."

The remainder of the book deals with various factors that are related to character. As to bodily aspects it is the more mobile tissues,—nerves, blood, muscle,—that are most significant although morphological types like Kretschmer's receive considerable emphasis.

The discussion of race is rambling but erudite. Morphologically race sets certain limits within which the individual constitution may vary considerably. It is quite similar with respect to character in that race sets limits within which it may vary.

The treatment of sex differences is rather conventional. There is no reference to experimental results except one questionnaire. The principal changes of character with age brought out are in the field of emotion.

Bodily chemistry is of somewhat doubtful influence on character. Some temperamental effects are attributed to arthritis, kidney disease, etc. The principal endocrine effects are included.

Pathological aspects of character include prenatal toxins, manias, hysteria, perversions. The author feels that there are some "born criminals" who lack "moral sentiment." Heredity is discussed at length. Character does not appear to follow Mendel's law. There are some traits such as nomadism that appear sex linked in inheritance.

The chapter on classification of character begins with a classification of classifications—philosophical, psychological, morphological. The author seems to lean toward Kretschmer. His own classification is threefold: isothymes have their dispositions properly proportioned; antisothymes have a disproportion of dispositions so that they are changing or cyclic; dyssynthymes are divided by incompatible rival dispositions. The final chapter on education points out the effect of education on certain dispositions.

The author has evidently read widely on the subject. He shows familiarity with the literature in many languages. He covers a lot of ground in detail and the average reader will probably peruse some portions hastily. The author goes back to the beginning of every topic so that he takes in much more than psychology. For instance the discussion of sex differences in character begins with the *X* chromosome

and its effect in determining sex. Consequently a good proportion of the book is not very closely related to the main problem. The reviewer is not disposed to criticize detailed points (except the author's enthusiasm for graphology) because the work is largely a presentation of what others have written. However it affords a good summary of the field and the attempts at classification are helpful. The point of view is philosophical rather than experimental. Few psychologists will care to read the book word for word but many of them will find portions that are worth their consideration.

HAROLD E. BURTT,
Ohio State University.

RUSSELL, CHARLES. *Classroom Scaler and Grader.* Ginn and Company, 1931.

This material consists of a manual together with a sort of slide-rule for finding various types of grades and scores. Recognizing that the usual grading of pupils in schools is very unsatisfactory, the author has constructed a device by which this grading can be made more uniform among teachers in general. The meaning of a grade is always in terms of the group to be graded. The author discusses various letter grade systems and indicates how they should be used. Six such systems are provided for use and others can be constructed. The value of M-Scale scores is emphasized and a scale for the rapid finding of these is included. Scales for finding achievement ratios and effort scores are also included in the material.

Undoubtedly the use of Russell's Classroom Scaler and Grader will lead to much more uniform grading among teachers who adopt it. For this reason it is strongly to be recommended, and also because it will reduce considerably the labor involved in the conscientious grading of students.

R. PINTNER,
Teachers College, Columbia University.

T. H. PEAR. *The Art of Study.* E. P. Dutton and Co., Inc., New York, 1931. 114 pages. Preface and Index.

This little volume, printed on heavy paper and attractively bound, is an elaboration of a series of broadcasts to school children on the subject of "How to Concentrate." It repeats with variations the maxims found in most books on this subject, but differs in that it presents the topics in an interesting manner, using a non-technical vocabulary and a friendly conversational style. Much of the book does not seem to bear

directly upon its title, *e.g.*, in chapter III is a discussion of individual differences in ease of learning, and in chapter X, Intelligence and Stupidity. But from the latter part of chapter IV, How to Listen, through chapter VIII, Practical Aids for Memorizing, there is plenty of meat. Chapter VI, How to Form Habits of Study, is worth the price of the book. The psychology is sound, it is worth any one's reading, and offers valuable suggestions to those who seriously want to form habits of study.

J. R. GENTRY,
Ohio University.

WARDEN, C. J. *Animal Motivation Studies: The Albino Rat.* Columbia University Press, 1931.

This book is a compilation of a series of investigations, most of which have appeared in different psychological journals and monographs within the past five or six years, carried out in the animal laboratory at Columbia University under the direction of Professor C. J. Warden. While most of the researches here reported are done by the collaborators, Messrs. T. N. Jenkins, J. H. Warner, E. L. Hamilton, H. W. Nissen and Marion Jenkins, the general introduction, the introductions to each report and Part VII, "A Comparison of Normal Drives," are written by Professor Warden.

Using the standardized "Columbia Obstruction Method" the following "drives" were examined under very carefully controlled conditions: hunger, thirst, sex (the last named in its many phases, such as the different incentive values of males, females, the latter in both oestrum and disoestrum stages; the effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on sex behavior), maternal, and exploratory drives.

While the "Obstruction Method" has been repeatedly presented elsewhere in the literature (also well described in this book) in connection with studies of animal "drives" by the Columbia investigators for the reader's benefit it will be briefly characterized here. A partial quotation from the book will suffice: "The box consists essentially of three compartments as follows: entrance compartment A, obstruction B, and incentive compartment C–D. . . ." It is while the animal is passing from compartment A through a tunnel in compartment B that he encounters an electric shock. The intensity of the shock resulting from current passing through the grid is determined by preliminary investigations with the white rat. In this way the investigators have reached optimal conditions of punishment for the animals with this technique. The measure of the strength of the drive isolated (partially at least)

for examination is in terms of the animal's *approaches, contacts* and the *actual crossings* made during his endeavor to leave the entrance compartment and to gain the incentive chamber. Instead of using "drive" in the sense that Dashiell and Holt do, these writers conceive it to be "an aroused reaction tendency which is characterized primarily by the fact that the activity of the organism is directed toward or away from some *specific incentive* such as food, water, animal of the opposite sex, etc." (pp. 14–15). The authors state specifically that "drive," as they use it, does not refer to a *physiological state* or *system*. Their conception of drive presupposes an incentive and a set of conditions imposed which determines the behavior of the animal. The reviewer suspects that there are many undomesticated animals which exhibit untold amount of behavior and never immediately reach the incentive; yet we have reasons to believe that a constant (more or less) source of internal stimulation keeps them reacting to their environment. After all then, it seems, that "drive" as Dashiell and Holt use it is not far wide of the mark. Then, too, Warden *et al.* rely upon the physiological or "tissue" conditions for explanations of the behavior patterns manifested by their animals. We might ask what term will we use for the "tissue" conditions which when deprived of certain substances initiate exploratory behavior in the animal? Somebody should examine the literature for the different usages of the terms, *reaction-tendencies, drives, motivation, mechanisms,* etc. These investigators have, at least, given their usage of *drive* and *incentive* an experimental meaning.

While the particular drives examined reached their peaks of intensity in the two sexes at different times the data were such that a ranking of the different "drives" could be made. Combining both male and female records the results show the order of rank from greatest to least, as measured by the maximum number of times the animals crossed the grid to get to the incentive, to be as follows: *maternal, thirst, hunger, sex,* and *exploratory.* These results, in many ways, are a bit out of line with some earlier results obtained by other investigators but with less rigidly controlled conditions.

Often the reader is conscious of the book being repetitious, certainly fewer introductions, in the reviewer's judgment, would be desirable; but as a ready reference to this series of studies on motivation it is good. In the Appendices are to be found four studies by Frances Holden, Warden and Hamilton, Hamilton, and Warden and Aylesworth in the order given. These problems are related to the major problems incorporated in the body of the book but they do not make use of the standardized "Obstruction Method," therefore the results obtained are not comparable to the other studies. Besides the description and experi-

mental analysis of standardizing the Columbia obstruction apparatus which constitutes Part I of this book, there are six other parts incorporating eight studies. These with the four studies in the Appendices make twelve studies in all compiled. At the end of each study is found a good bibliography on each study. The book as a whole is a critical analysis of methods used in obtaining comparable data on motivation in animal behavior.

JAMES R. PATRICK,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED*

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

Australian Council for Education Research. First Annual Report, 1930–31. 145 Collins Street, Melbourne, Victoria. 46 pp.

Building Personality in Children. CARRY MYERS. Greenberg, Publishers, 160 Fifth Avenue, New York City. 1931. $2.50. 360 pp.

Cosmic Consciousness: A Study in the Evolution of the Human Mind. RICHARD MAURICE BUCKE. E. P. Dutton & Company, New York City. 1931. $5.00. 385 pp.

Educational Psychology. WILLIAM CLARK TROW. Houghton Mifflin Company, New York City. 1931. 504 pp.

Marie Louise, Napoleon's Nemesis. J. ALEXANDER MAHAN. Thomas Y. Crowell Company, New York City. 1931. $3.75. 364 pp.

The Measurement of Interests. DOUGLAS FRYER. Henry Holt & Company, New York City. 1931. $4.50. Student edition, $3.60. 488 pp.

Mental Measurement of Pre-school Children. RACHEL STUTSMAN. World Book Company, Yonkers, N. Y., 1931. $2.20. 368 pp.

Ourselves and the World. F. E. LUMLEY and BOYD H. BODE. McGraw-Hill Book Company, New York City, 1931. $2.00. 591 pp.

Readings in Experimental Psychology. WILLARD LEE VALENTINE, Editor. Harper and Brothers, New York City. 606 pp.

Speech Pathology. LEE EDWARD TRAVIS. D. Appleton & Company, New York City, 1931. $4.00. 331 pp.

A Study of the Unsupervised Behavior of a Group of Institutional Children. MAPHEUS SMITH. Marshall and Bruce Company, Nashville, Tenn., 1931. 40 pp.

Work-Book in Abnormal Psychology, Form A. JOHN J. B. MORGAN. Longmans, Green & Company, New York City. 1931. 88 pp.

A Work Book in Educational Psychology. HARVEY C. LEHMAN and STUART M. STOKE. Ohio University Book Store, Athens, Ohio, 1931. $1.35. 75 pp.

* Mention here does not preclude further comment.

Directory of American Psychological Periodicals

(Continued)

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 500–600 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Philosophy—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription. $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology and The Pedagogical Seminary—Worcester, Mass.: Clark University Press. Subscription $7. 600 pages annually. Edited by Carl Murchison. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300-400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $5. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (5 Numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Lancaster: The Science Press Printing Co. Subscription $3.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 500 pages each. Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

(*Continued on Third Cover*)